

REVIEW ARTICLE

Constraints on Galaxy Structure and Evolution from the Light of Nearby Systems

Gerald Cecil and James A Rose

Dept. of Physics & Astronomy, University of North Carolina at Chapel Hill, NC 27599-3255
USA

E-mail: cecil@physics.unc.edu, jim@physics.unc.edu

Abstract. We review knowledge of galaxy structures obtained by their emitted light and in the local universe where they can be studied in great detail. We discuss the shapes of, and stellar motions within, galaxies, compositional clues derived from their spectra, and what luminous matter implies about their dark matter content. Implications on the current theory of hierarchical galaxy formation are explored.

1. Introduction

We have entered the age of ‘precision cosmology’. Careful measurements of how objects of known radiant power have dimmed with distance, the angular size of known distances at early epochs, and the power spectrum of the matter distribution at the present epoch, have led to a remarkably accurate cosmological model. That is, the expansion history of the universe, which is governed by the density of visible and dark matter (Freeman and McNamara 2006) and by dark energy, has been well constrained. The tools of observational cosmology are amazingly diverse. Type Ia supernovae (SNe) — the deflagration/detonation of a white dwarf star after it accretes sufficient mass to exceed the pressure support of degenerate electrons — provides a high-luminosity ‘standard candle’ for mapping the dimming of sources over cosmological distances, constraining directly how the Hubble expansion parameter changes with cosmic epoch. Detailed maps of the power spectrum in angular scales of temperature variations of the cosmic microwave background radiation reveal various harmonics of density fluctuations at the time when the universe had cooled sufficiently for electrons and protons to ‘recombine’ into hydrogen atoms. The scale of these fluctuations depends on the matter energy density in the early universe, and the angular scale observed at the present epoch depends on the expansion history since this recombination. Hence, the temperature maps constrain the cosmological world model. Finally, the power spectrum of the 3D distribution of matter on large scales, as traced by luminous galaxies, also constrains both the spectrum of the initial density fluctuations and the combination of matter and dark energy that governs their evolution.

The recent success of observational cosmology poses both new challenges and opportunities. Perhaps most exciting is to quantify structure growth. In the favoured dark energy plus cold, dark matter (Λ -CDM) model, non-baryonic dark matter (DM) is non-relativistic long before the epoch where neutral hydrogen atoms (HI in astronomical jargon) first form. From the minimum (Jeans) mass that can collapse by self gravity at that time ($\gtrsim 10^6$ solar mass, M_\odot), cold dark matter (CDM) clumps in a hierarchy of increasing mass.

This scenario has been simulated extensively into the non-linear regime, but is largely untested by observations. Temperature anisotropies in the cosmic microwave background radiation probe only the *start* of the linear regime of structure formation as gravity drags on cosmic expansion. While anisotropies support the idea that the seeds of non-linear structure have Gaussian amplitude spectrum with random phases (i.e., are spatially uncorrelated), unclear is how coherent structure evolves into the non-linear regime, and if the radiant power (luminosity in astrophysical jargon) and mass scales and morphologies of the observed structures are consistent with the CDM hierarchy.

To document how structure evolves from when galaxies first appeared at redshift $z \sim 5$ until the present epoch it is essential to understand the evolution of visible baryons as density-biased tracers of the underlying DM. Even studies that map DM more directly through gravitational lensing must know the structure of the lens by its visible material. Thus, all studies of structure evolution depend on a better understanding of galaxies.

Determination of our cosmological world model offers an exciting opportunity for students of galaxy evolution. Prior interpretation of, for example, number counts of galaxies as a function of redshift, depended on both galaxy evolution and an uncertain world model. An accurate model now decouples geometry from galaxy evolution observed over a significant redshift, hence lookback time.

Advanced ground- and space-based telescope and detector technologies from x-ray to millimetre wavelengths allow study of galaxy evolution *in situ* at high redshift. To mention only a few examples over the past decade: the development of ‘optical’ telescopes of up to 10.5-meter aperture that are coupled to multi-slit/fibre spectrographs, advanced detectors at infrared (IR) wavelengths, the advent of imaging at sub-millimetre wavelengths, high-spatial resolution UV to near-IR images from the Hubble Space Telescope (HST), detailed studies by x-ray observatories of hot plasma in the Milky Way galaxy (MWg) interstellar medium (ISM) and in the intracluster/group medium, and mid-IR views from the Spitzer Space Telescope.

While there has thus been a stampede toward objects at high redshift, to understand these challenging, photon-starved measurements require multi-waveband constraints on comparable objects nearby. We must assess how baryons have concentrated into galaxies through, among other processes, star formation (SF) and evolution, nucleosynthesis of heavy elements, and dynamical processing of stars and dissipational ISM. Because galaxies are proximate enough to interact strongly by tides, we must determine how these processes are affected by environment (rich galaxy cluster versus small group).

For those oriented toward cosmology, understanding galaxy evolution is a necessary prerequisite to using galaxies in bulk to trace the evolution of the dominant DM and to constrain the mysterious dark energy. Others will research the diverse and fascinating physics of galaxy evolution. Our review is aimed at those who are preparing to follow either path. We assess what properties of galaxies have been established recently in the *nearby* universe ($z \lesssim 0.1$, an insignificant lookback time compared to the age of the universe). In number, 80% are gas-rich spiral galaxies (Sgs) and smaller dwarf irregular galaxies (dIrr’s), the rest are gas-poor elliptical (Eg) and S0 (called collectively, early-type) galaxies and dwarf spheroidal galaxies (dSph’s). We address questions about galaxy structure and evolution using images and spectra of much higher intrinsic spatial resolution and signal-to-noise ratio (SNR) than can be obtained on high redshift counterparts. Our theme throughout is the perhaps surprising degree of uncertainty that prevails over most topics in galaxy structure and evolution.

Section 2 introduces modern observations of stars and DM. At the top level, we seek the main driver(s) of galaxy evolution. Does a single process dominate? Our main observables are the details within the bimodal separation of galaxies into luminous ‘red and dead’ early-type spheroids and blue star-forming discs. How are these connected? Are Egs a single family

differing only in mass? Section 3 examines their 3D structure for universal characteristics. A consistent interpretation of their surface photometry and kinematics requires the introduction of many dynamical assumptions. Section 3.5 discusses their uncertain DM content. Section 4 examines the apparently similar spheroidal bulges of Sgs, including those of our MWg and our neighbour Sg Andromeda (M31). Are they primordial, were they built from mergers, or did they arise from secular (i.e., long-term) processes driven by instabilities?

In section 5 we study Sg discs of various thickness. Are discs stable and axisymmetric? DM halos affect stability, so we address secular effects in section 6 and consider radial matter transfer. Are discs distinguished only by their level of SF? We consider the multi-component ISM of gas, dusty soot, magnetic field, and cosmic rays. At the bottom of the potential well is often a supermassive black hole, best studied by its induced stellar motions.

In section 7 we consider interactions of stars with the hot gas and with other galaxies in rich clusters. Do galaxies switch between spiral and ellipsoidal forms? How has environment modified SF? We preview some new instrument capabilities in section 8, and summarize in the last section. We assume that the reader grasps the basics of the current cosmological world model (for example the textbook by Ryden 2002), but we minimize astronomical jargon and suggest review articles as broader topics emerge. Appendix A reviews key astrophysical processes to understand before considering galaxy evolution; Table A.1 lists acronyms used throughout.

2. What do we learn from starlight?

We first summarize observables from galaxy starlight. Limiting uncertainties are *systematic* (external), not statistical (internal), in most astronomical datasets of relevance to galaxy evolution because most dynamical parameters are not observed directly. Instead, we shall see that they arise from correlations between, and calibrations from, often heterogeneous photometric and kinematical datasets whose inter-transformations introduce systematics. Uncertain zero points and incompleteness biases that weigh points in a correlation can be further confounded by the inadvertent inclusion into the study sample of unrelated less luminous foreground or more luminous background galaxies; the large range of galaxy luminosities and the complications of patchy obscuration by dust mean that reliable distances are a prerequisite to minimising such bias. Differential measurements can minimize systematics. Computer simulations to quantify errors are necessary in observational astrophysics, complicated by both the large and very small sizes of astronomical datasets.

2.1. The Milky Way galaxy

Observing stars over a wide range of masses and evolutionary states, one seeks to correlate their dynamical, SF, and chemical enrichment histories with 3D space velocities, ages, and chemical compositions. Stars live mostly in the relatively low luminosity phase of core hydrogen-to-helium fusion. They spend the last $\sim 20\%$ of their lives in a series of higher luminosity phases. High-mass stars have relatively brief lives compared to low mass ones.

Only in the MWg can we obtain such extensive data on individual, especially lower mass unevolved, stars. But only a minuscule fraction of its $\sim 10^{11}$ stars can be so detailed from our vantage point 8.0 ± 0.4 kpc (Eisenhauer *et al* 2003) off-centre. Indeed, we can approximate this program only within the ~ 50 pc radius Solar Neighbourhood where we can obtain accurate stellar distances by trigonometric parallax from opposing points on the Earth's orbit. Otherwise, use of secondary measures of distance built upon parallaxes amplifies uncertainties. Over the last decade, parallaxes and photometry from the Hipparcos satellite

above atmospheric blur has been combined with ground-based spectra and photometry to calibrate the parameters of many nearby stars accurately (Lebreton 2001); future astrometric space missions aim to expand this sphere a hundred-fold in radius. Note that many stars with excellent parallax measurements have relatively poor spectroscopy and even photometry; these deficiencies are being rectified by ongoing surveys. Finally, our view of the MWg in visible and UV light — where the most powerful diagnostics lie — is obscured by interstellar dust. Only outside this waveband does our view clarify along many sightlines.

Now consider velocity vectors. With parallax in hand and a radial (line of sight, l.o.s.) velocity measured from the Doppler shift of narrow absorption lines in the stellar photosphere, we must obtain the star's tangential motion across the sky by measuring at two widely spaced times its sky position relative to more distant objects that define a fundamental reference frame. As detailed in Appendix A.2, the orbit inferred from the gravitational potential is the essential dynamical input in the study of galaxy evolution. Key too are age and chemical composition, which are extracted from spectral analysis as summarized in Appendix A.1.

2.2. Local Group galaxies

The MWg helps to bind the Local Group of galaxies. Across ~ 2.5 million parsecs (Mpc), two large Sgs — the MWg and larger M31 — dominate ~ 25 satellites; a third smaller Sg M33 ($\sim 10\%$ of the mass of the MWg) is associated with M31. Together, the three Sgs emit 90% of the light of the Local Group. With M33, the other two large satellites of M31 are the dwarf Eg M32 and dwarf S0 NGC 205. The largest satellites of the MWg are the barred Large and irregular Small Magellanic Clouds (LMC and SMC). Local Group members are still being discovered, especially near the dust obscured plane of the MWg.

Local Group members are close enough for some resolution into individual stars, especially with HST. Luminous stars can be studied individually, tending to be either young and massive or in later evolutionary stages. The motions of the Magellanic Clouds across the sky have been measured from stars (Kallivayalil *et al* 2006a, Kallivayalil, van der Marel and Alcock 2006b). The near side of the LMC is 10% closer than the far. The resulting 20% brightness difference of stars of the same stellar type gives a rudimentary sense of the LMC's 3D structure. All other Local Group members are either too small or distant to yield reliable information along the l.o.s. Hence the MWg, and to a far lesser extent the LMC, are the only galaxies for which 3D information is obtainable directly.

High-resolution spectra, from which chemical abundances can be derived, have been obtained for individual stars in some MWg satellites. Abundances constrain the chemical history of a galaxy. For example, in section 3.6.5 we compare abundances in the Sagittarius dIrr to those in the thick disc and halo of the MWg, to constrain the assembly of the MWg.

2.3. Other 'nearby' galaxies

Individual stars outside the Local Group are virtually unmeasurable. Instead, the 3D structure of a galaxy is extracted from starlight projected onto the sky and integrated spatially across each image pixel. One must rely on the integrated l.o.s. velocity field (velocity function) that is broadened by dispersed stellar velocities. In addition to the pixel's spectral content and mean light flux, variations in the surface brightness (SB) that arise from Poisson fluctuations in the number of bright, evolved stars within each pixel can be used to constrain the galaxy distance (Tonry and Schneider 1988).

Most galaxies have only a cosmic redshift-implied distance; only $\sim 10,000$ have distances determined independently with accuracy sufficient to separate a galaxy's l.o.s.

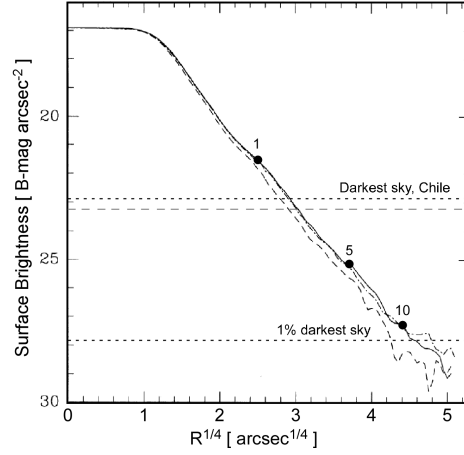


Figure 1. SB of the nearby Eg NGC 3379 (photometry from Capacciolo *et al* 1990) with ● at radii xR_e for $x = (1, 5, 10)$, the darkest night sky at a Chilean site (top), and its 1% level (lower), and the value at near-Earth orbit (— · —). Because SB depends empirically on radius as the $1/4$ power, this variable is plotted on the ordinate to highlight deviations; profiles along both the apparent major and the minor axis are shown. The galaxy profile within ~ 1.5 arcsec radius is flattened by the blurring effects of the Earth's atmosphere.

peculiar velocity due to gravity from the bulk cosmic expansion redshift. In short, we know at best only three of the six spatio-kinematical descriptors of galaxy structure. In situations of high symmetry, we can invert to constrain the spatial variation throughout some of the volume. As discussed in section 3.3, asymmetries in absorption line profiles can constrain the structure of less symmetrical systems.

2.4. State of the art

At its average SB ($V = 22.5$), the moonless night sky over angular area 1 arc-second² emits the radiant flux of a giant star at 0.2 Mpc or a Sun-like dwarf star at 32 kpc. Adaptive optics systems on telescopes reduce the area of the effective detection box, hence sky flux, at least another ten-fold, increasing the detection distance of a single star at constant SNR more than three-fold. From the ground, the SNR increases at best as the square-root of the increased exposure time. HST goes considerably deeper than ground-based facilities somewhat because of a $\sim 17\%$ on average reduced V-band background (which is still variable because of unpredictable, cloud dependent Earthshine) but mostly because its diffraction limited image tightens the detection box to $\sim 0.02(\lambda/500\text{nm})^2$ arc-second².

The radial light profiles of spheroidal galaxies projected onto the plane of the sky are characterized by an ‘effective radius’ R_e within which half of the galaxy luminosity lies. Figure 1 compares the SB of the moonless night-sky to those of spheroidal galaxies. With diligent sky-subtraction, accurate photometry can be obtained beyond $10R_e$. Slit spectra can be obtained to $5R_e$, and micro-lens- or fibre-coupled spectrographs (field segmentation techniques whose optics tend to scatter more light than open slits) can be effective out to $1 - 2R_e$. Until recently, selection effects excluded galaxies of considerable angular extent but of constantly low SB; section 5.4.4 discusses this population.

A complication, becoming historical, to transforming insights between the local and distant universe is that we usually observe different rest wavelengths at high redshift than we

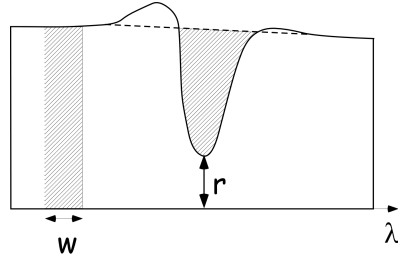


Figure 2. Defining equivalent width w and residual intensity r of an absorption line spectrum. The vertical axis is radiant flux, zero at bottom. A ratio of residual intensities defines an index. To define w , the area notched from the slanted stellar continuum is repeated at left (or right). As shown, equivalent width depends even in data of infinite SNR on how one draws ‘the’ continuum across a line and where the two sidebands are defined; this is especially true at blue wavelengths where lines crowd. For an index based on r , the continuum level is irrelevant but scattered light must be controlled. $\sim 45\%$ of Egs show weak emission lines, which must be excised for a pure absorption spectrum.

do nearby. Only recently have UV-optimized spacecraft (for example, the Galaxy Evolution Explorer [GALEX] and Far Ultraviolet Spectral Explorer [FUSE]) and near-IR ground- and space-based spectrometers provided images and spectra with sufficient wavelength coverage to transform accurately. Another complication is the inadvertent inclusion of luminous background objects, often star burst and/or active galactic nuclei (AGN), whose bright regions are uncharacteristic of the underlying stellar mass. These ubiquitous contaminants increase with survey volume and complicate the connection to nearby objects.

Signatures of chemical abundances concentrate at wavelengths shorter than $\lambda 550$ nm (see Figure 2 of Bland-Hawthorn and Freeman 2004; for example), the U, B, and V filter bands. One needs $\sim \lambda 0.01$ nm resolution ($R = \lambda/\Delta\lambda \sim 40,000$) to measure accurately many elements mostly by minimising systematic errors from otherwise blended lines. Stellar chemical abundances are determined by high-dispersion spectroscopy of many, sometimes weak absorption lines from key elements at various stages of ionization. This program is feasible on even halo MWg stars with the largest telescopes and most efficient spectrographs; an error of $\lambda 10^{-4}$ nm in the equivalent width (defined in Figure 2) of an absorption line yields an abundance error of 0.02 dex. F-K stars emit most of their energy in the U- to L- ($\lambda 5000$ nm) bands that are all accessible from the ground. A SNR > 10 measures [Fe/H] with errors of ± 0.1 dex and errors in radial velocities of ± 2 km s $^{-1}$. \ddagger A six hour exposure with the latest spectrometers on the Keck telescopes measures the intensities of H β absorption lines in even dwarf Virgo cluster galaxies (~ 16 Mpc) to $\sim \lambda 0.02$ nm equivalent width accuracy. In section 3.6.1 we show that this translates into an age uncertainty of ~ 4 Gyr for a 12 Gyr-old star cluster; uncertainties are exponentially smaller for younger clusters.

Over the past decade, several wide-field surveys were made of the MWg and other galaxies. Notable in the MWg are microlensing experiments (section 5.3.2) and the Wisconsin H α Mapper (WHAM, Hafner *et al* 2003) survey of the ionized ISM. Extragalactic surveys include the Sloan Digital Sky Survey (SDSS, York *et al* 2000), the Two-Degree Field Galaxy Survey (Colless *et al* 2001), and the near-IR 2MASS (Struck *et al* 2006) and Deep Near Infrared of the Southern Sky surveys (DENIS, Epchtein *et al* 1994) that dip closer to the plane of the MWg to find galaxies that are mostly dust obscured at shorter wavelengths.

\ddagger A chemical abundance in brackets is measured in units of dex, i.e. $[A/X] \equiv \log_{10}(N_A/N_X) - \log_{10}(N_A/N_X)_\odot$, where N_A and N_X are the number density abundances of elements A and X. Positive dex means more metals than in the Sun.

3. Ellipsoidal Systems

Egs (and the bulges of disc galaxies), by their simple SB profiles and lack of an extensive ISM and ongoing SF, would seem to be the most straightforward stellar systems hence the natural starting point to study galaxy evolution. To focus discussion, we first frame basic questions within the context of the evolution of structure in the universe:

- (i) Can the true 3D shapes of Egs and disc galaxy bulges be recovered uniquely from their 2D SB? Are they primarily oblate, prolate, or perhaps more complex (triaxial) spheroids?
- (ii) Do Egs form a single homologous family (hence formation mode) or fundamentally different families (formation modes) over their entire million-fold mass range?
- (iii) Did SF/chemical evolution complete in only a few dynamical times (< 1 Gyr), or is there evidence for extended (perhaps discontinuous) SF?
- (iv) Are Egs dominated by DM, and how is DM spread compared to luminous matter?

We summarize work that has sought to understand SF and chemical evolution histories in diverse environments and as functions of galaxy luminosity (and mass) to test homology, to establish the distribution of their intrinsic shapes, to learn the dynamical rôle of a central supermassive black-hole, and to constrain their DM content.

3.1. Current observational capabilities

Before addressing the observed nature of Egs, it is useful to establish the capabilities of current large (6.5-meter+ aperture) ground-based telescopes with efficient visible-band/near-IR detector mosaics and multi-aperture spectrographs, and of rapidly expanding spectroscopic surveys. What are the relative merits of large angular-area, high-SNR, spatially-resolved data of a single (or a few) nearby object(s), as opposed to more modest data on huge numbers of galaxies in surveys such as the SDSS? The datasets deliver complementary insights, but it is instructive to examine for specific questions whether exhaustive data on one object are more illuminating than less extensive data on many. One can often recover the distribution in properties of a large sample from statistical analyses of the distribution of key observational parameters by imposing Bayesian prior information. The biggest challenge for the large sample approach are selection biases that occur when working with flux-limited samples of spatially extended but diffuse objects. The problem is more severe when comparing distant to nearby samples, where it is generally uncertain how the luminous objects detected at high redshift relate to the bulk of galaxies at the present epoch.

During the 1980's and 1990's, long-slit spectrographs on 4-meter aperture telescopes measured Eg kinematics, statistics improving with increased CCD sensitivities and pixels. As noted in section 2.4, by R_e in a massive Eg, the SB of B-bandpass starlight has dropped typically to that of the moonless night sky. By $2R_e$, one is working ~ 6 times fainter than the sky SB. With a 4-meter aperture telescope, one obtains spectra typically out to $\sim 1.5R_e$ in reasonable exposures. To constrain the internal structure of the galaxy requires spectra at multiple position angles (PA's) on the sky, at least along the photometric major and minor axes. In contrast, photometric profiles for Egs could be traced down to 1% of the sky SB with $\pm 30\%$ systematic uncertainty from the removal of its time-variable OH emission-line glow. Another limiting factor is often the care taken in acquiring 'flat field' images to calibrate spatial variations in photometric sensitivity across the detector and through the optics. Particularly damaging is scattered light within the instrument+telescope.

Three major advances have increased sensitivity substantially.

- (i) Spectroscopic studies have progressed from sparse, long-slits along select PA's reaching $\sim 2R_e$ (for example Bak and Statler 2000) to the full spatial coverage over $\sim R_e$ of the SAURON integral-field spectrograph (IFS) (de Zeeuw *et al* 2002) on > 70 early-type galaxies in both clusters and the field. SAURON constrains within R_e , a complicated region containing counter-rotating stars, warps, and other transient structures. Its optics deliver > 1500 contiguous spectra per exposure in either a lower spatial resolution mode used to maximize spectral SNR at large radii, or a higher resolution mode used on brighter regions near the galaxy core. While mapping kinematics to 6 km s^{-1} accuracy, its spectra also yield fluxes of some spectral indices to $\lambda 0.01 \text{ nm}$ accuracy in equivalent width. We introduce SAURON results below, as appropriate. Other smaller-field IFS's (some using an adaptive optics system to sharpen ground-based images by nulling part of the effects of atmospheric turbulence) have detailed stellar cores.
- (ii) Development of the 'nod-and-shuffle' technique of coordinated telescope and detector shifts (Glazebrook and Bland-Hawthorn 2001) to subtract night-sky OH emission spectral bands to Poisson statistical limits. Implemented on 8-meter+ aperture telescopes with efficient spectrographs, this technique cancels many systematic errors to obtain spectra at much fainter light levels than previously. Note that long-slit spectroscopy is an inefficient data collection tactic, because it covers outer radii only sparsely where large areas are needed to boost SNR. With an IFS, coverage is continuous over the field of view, with outer regions receiving the higher weight necessary to boost SNR. Nevertheless, because of the purity of sky subtraction through an open long-slit plus nod-and-shuffle, the sparse coverage at very low SB of long-slit spectra still complements the full coverage at higher SB of an IFS whose field-segmenting optics scatter more light.
- (iii) The advent of wide-angle imaging and spectroscopic surveys. Specifically, the 2dF survey on the 3.9-meter Anglo-Australian Telescope produced $\sim 250,000$ galaxy redshifts, and the SDSS on a 2.5-meter telescope produced multi-bandpass photometry of $> 10^7$ galaxies and multi-fibre spectra of $\sim 10^6$ galaxies. With proliferating multiple-CCD cameras and dedicated robotic, wide-field telescopes for executing all-hemisphere imaging and spectroscopy, the future of large scale surveys is extremely promising.

3.2. Intrinsic 3D shape

We begin discussion§ of Egs with a deceptively simple question: what is their intrinsic shape? Appendix A.2.7 noted that obtaining the 3D distribution from the Abel inversion of the 2D SB distribution is ill-conditioned, being highly unstable to details of noise and data artifacts; the inversion can only be achieved with an assumed 3D symmetry. Thus, even very high quality SB data on the nearby, nearly circularly symmetric 'standard' Eg NGC 3379, yield no clear answer on whether its slight ellipticity (Figure 1) implies oblateness, prolateness, or even triaxiality. It may well harbour a weak, nearly face-on disc (Capaccioli *et al* 1991).

Eg shapes were therefore first constrained in galaxy samples by trending the central SB against their degree of ellipticity. galaxies of different ellipticity. to map central SB versus degree of ellipticity. The seemingly well-posed hypothesis that prolate galaxies viewed edge-on should have higher SB than edge-on oblates has no clear-cut answer, although the tendency has been to favour oblates (for example Marchant and Olson 1979). However, this statistical approach trades the uncertainty of deprojecting a SB profile to 3D to that of extracting the true parent distribution of 3D shapes from an observed distribution of peak SB's and ellipticities,

§ Readers who are unfamiliar with gravitational stellar dynamics should first read the material in Appendix A.2.

assuming of course that there *is* a well-defined parent distribution. In fact, Lambas *et al* (1992) conclude that Egs are neither strictly oblate nor prolate.

Twisted isophotes whose ellipticity and PA of projected major axis change with SB occur in many galaxies. Section 4.2.3 of Binney and Merrifield (1998) shows that twisted isophotes originate naturally when we view misaligned a triaxial system whose principal axes are aligned and orthogonal at all intensities while varying in axis ratio systematically with SB. But twists also result from strong tides between galaxies, and indeed the most extreme ones are found in close galaxy pairs (di Tullio 1979, Kormendy 1982). Add the aforementioned possible disc lurking in some Egs, and 3D shapes cannot be determined.

In summary, a seemingly simple question about the supposedly simplest galaxies leads to a surprisingly unsatisfactory result. To progress, one must investigate kinematics.

3.3. Internal motions

By the mid-1990's it was clear that most luminous Egs have isophotes whose major axial PA's twist with radius, implying triaxial shapes. Because triaxial shape implies a large fraction of stellar orbits with additional isolating integral(s) (see Appendix A.2.3) beyond E and L_z (Schwarzschild 1979), the distribution function f will have anisotropic velocity dispersions. A signature of a rotating system with isophotes twisted by projected triaxiality is rotation along *both* projected axes, which is seen in many Egs (for example Schechter and Gunn 1978). Evidence for anisotropic velocity dispersions is more subtle. Any such system will flatten more than could arise from the observed rotation. This qualitative global statement lends itself naturally to quantification by the tensor virial theorem (see Appendix A.2.5), as developed by Binney (1978), further discussed in section 4.3 of BT, and reexamined by Binney (2005) in the light of the new capabilities of IFS's.

The main indication of a triaxial shape is misaligned rotation, with motion from tube orbits (Appendix A.2.4) streaming along the true minor-axis, but observed along the *projected* minor axis of an apparent axisymmetric galaxy. Asymmetric absorption-line profiles are quantified with up to 6th-order Gauss-Hermite velocity moments. Combining absorption-line long-slit spectra to map rotation speeds and surface photometry to map shapes, one interprets with Jeans models (first two moments) from the tensor Virial Theorem or more general three-integral models (higher moments too) to constrain the internal motions of spheroids.

In an axisymmetric galaxy viewed edge-on that rotates about the z-axis, the potential energy tensor

$$W_{xx} = W_{yy}; W_{ij} = 0 \ (i \neq j) \quad (3.1)$$

and similarly for tensors Π and T (random and ordered motion); in addition, $T_{zz} = 0$. The tensor Virial Theorem then gives

$$\frac{2T_{xx} + \Pi_{xx}}{\Pi_{zz}} = \frac{W_{xx}}{W_{zz}} \quad (3.2)$$

If we define

$$2T_{xx} = \frac{1}{2} \int \rho \bar{v}_\phi^2 d^3x = \frac{1}{2} M v_0^2 \quad (3.3)$$

with M the total mass, v_0^2 the mass-weighted mean-square rotational part of the velocity function, and σ_0^2 the mass-weighted mean square random part of the velocity function, we obtain

$$\frac{v_0^2}{\sigma_0^2} = 2(1 - \delta) \frac{W_{xx}}{W_{zz}} - 2 \quad (3.4)$$

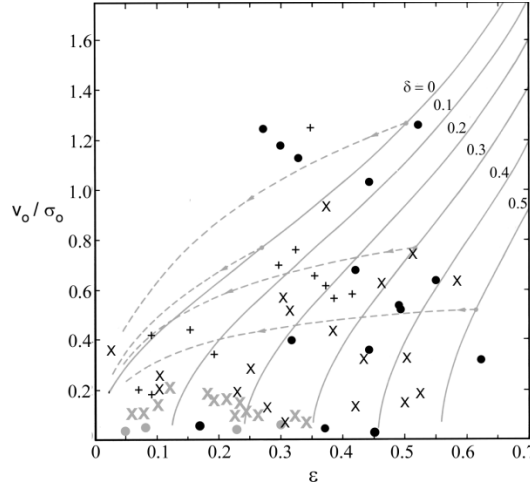


Figure 3. How observables (ratio of maximum rotational to mean random motions V_0/σ_0 and photometric ellipticity $\epsilon = 1 - \text{axial ratio}$) averaged over an oblate spheroidal Eg are related for various values of the velocity anisotropy δ —. Lines (---) map to the left the theoretical δ curves drawn at right as our viewing angle is varied from edge-on (right) to face on (lower left). Plotted are light-weighted averages across individual luminous Egs (\times from Illingworth 1977, Bertola and Capaccioli 1975) that mostly show no rotational support, and Virgo dwarf Egs (\bullet from Geha *et al* 2002 and van Zee *et al* 2004) and lower luminosity Egs ($+$ from Davies *et al* 1983) that mostly do; grey denotes an upper limit of the measured V_0/σ_0 .

with $0 \leq \delta \leq 1$ measuring the anisotropy of the velocity dispersion tensor (i.e., $\Pi_{zz} = (1 - \delta)\Pi_{xx}$). If the isodensity surfaces are ellipsoids, then W_{xx}/W_{zz} depends only on their average photometric ellipticity ϵ and not on their density profile $\rho(r)$ along galaxy radii. If v_0 and σ_0 can also be estimated by averaging spectra, then — lines in Figure 3 plot the relation between the spatially averaged v_0/σ_0 , ellipticity ϵ , and anisotropy parameter δ .

There are two complications. First, (3.4) assumed that we view the galaxy edge-on to its true major axis. Otherwise the observed eccentricity and velocity function alter. The dashed lines in Figure 3 show different δ hence viewing angle i that is usually unknown, so δ is degenerate. However, different parts of the diagram indicate clear oblate rotation or anisotropic flattening. Second, although the tensor Virial Theorem is a global average of observables, most of its results arise from a few long-slit spectra through the nucleus along the major axis (and minor axis and crossed at 45° when more spectra are obtained). Spectra normally do not extend beyond R_e , so are not global averages. An assumption must convert from observables such as maximum rotational velocity and central velocity dispersion to global values required by the tensor Virial Theorem.

Despite these deficiencies, our understanding of Eg dynamics is based on this Theorem. Early work (Illingworth 1977, Bertola and Capaccioli 1975) found that most luminous Egs flatten largely by velocity anisotropy, not rotation. Subsequent work (Davies *et al* 1983) led to the surprise that lower luminosity dwarf Egs are flattened mostly by rotation. Recent attention has focused on Virgo cluster dwarfs (Geha *et al* 2002, van Zee *et al* 2004). While most do not rotate, some do so fast enough to flatten. Figure 3 summarizes these results.

The tensor Virial Theorem gives the global balance between the inferred rotation and velocity dispersion anisotropy (hence further infers triaxiality and the potential), at least when we have a reliable inclination. Details on the internal structure of Egs can come only by fitting to some dynamical model the spatially resolved spectra of both streaming and random motions

over an area of the galaxy. Statler and collaborators have fit spectra mostly from Davies and Birkinshaw (1988) out to $\sim 1.4 R_e$ using Stäckel potential models with minimal prior assumptions about dynamics. The resulting bivariate probability distribution of flattening c_L and degree of triaxiality T often yields an oblate best fit for an individual galaxy, but the result is usually not well constrained. Instead, Bak and Statler (2000) fit to the parent shape distribution of Davies and Birkinshaw (1988)’s 13 Egs, using the same Bayesian models. They seek the shape distribution of the entire family, which can be estimated more robustly than the shape of any single galaxy. Figure 4 contours the intrinsic shape of their sample as a bivariate likelihood distribution in T and c_L . At left is the distribution from spectra and surface photometry, with both prolates and oblates present but with oblates more likely and strong triaxiality rare. In contrast, the right panel shows the parent distribution if only photometry is used. While flattening is still well constrained, triaxiality is not.

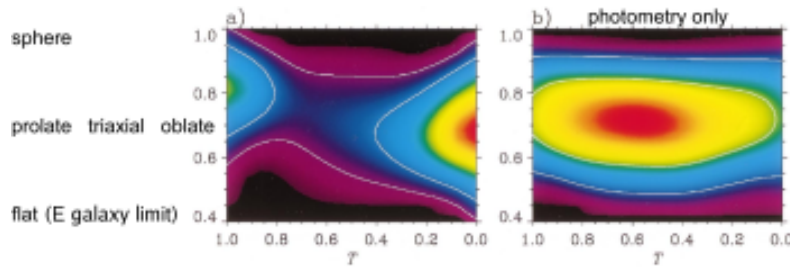


Figure 4. (From Bak and Statler 2000; used with permission.) The parent distribution of intrinsic shapes of a sample of 13 Egs with high-quality, multiple position-angle slit spectra, plotted in the space of triaxiality T and flattening c_L of the light distribution. Contours enclose 68% and 95% probabilities under the assumption that the parent population has no preferred orientation to our l.o.s. a) Using photometric *and* kinematic data, a bimodal distribution is evident, with more probability of oblate than prolate objects. The probability of them having very triaxial shape is low. b) Photometric data alone constrain only the flattening.

In summary, in proceeding from analyses of individual galaxies (e.g., NGC 3379) to a set drawn from a distribution, shapes are better constrained but at the expense of added dynamical assumptions on the parent distribution. Bak and Statler (2000) find that if rotation is assumed to be mainly ‘disc-like’ (decreasing rapidly from the major axis), then triaxiality is indeed rare. However, if rotation is assumed to have ‘spheroid-like’ constant rotation speed at all latitudes, then triaxiality is common. Results are still hostage to model assumptions.

The advent of IFS spectra that map the velocity function across the inner $\sim R_e$ now allow application of the tensor Virial Theorem to a representative sample of Egs with fewer assumptions (Binney 2005). For example, Verolme *et al* (2002) use SAURON IFS spectra to constrain the shape of M32. Assuming axisymmetry, they obtain intrinsic flattening 0.68 ± 0.03 . Krajnović *et al* (2005) is another illustration of modern orbit fitting. Cappellari *et al* (2005) is a preliminary analysis of part of the SAURON sample, in the spirit of Figure 3. They show that those galaxies divide into weakly triaxial and nearly isotropic slow-rotators, and nearly oblate and anisotropic fast-rotators.

The reader may wonder if this discussion has led to a result. Clearly, the CBE has not been solved uniquely. That results (for example on whether a sample of Egs is mainly prolate or oblate) are as varied as the underlying assumptions shows that we are still far from uncovering shapes and establishing the dynamics of Egs, especially using only photometry. Yet, significant progress *has* been made: the tensor Virial Theorem, supported by isophotal twists evident in some Egs, shows conclusively that the more massive Egs are flattened mainly

by velocity dispersion anisotropy, while intermediate mass Egs tend to be oblate rotators. This is a key constraint on scenarios of galaxy evolution. Further progress should come soon as IFS samples are analysed with statistical techniques such as those of Bak and Statler (2000); for example, see Statler *et al* (2004).

3.4. Global scaling parameters

Recall a question posed at the start of this Section: are Egs a single homologous family over their full mass range? While some investigators consider the range $\sim 10^6$ to $\sim 10^{12} M_\odot$, we adopt the conservative approach that Eg luminosities range between $-18 > M_B > -22$, dwarf Egs range between $-13 > M_B > -18$, and galaxies still fainter fall into the dSph and/or globular cluster category. While the connection of dSph's (studied extensively in the Local Group) to the main family of d/Egs is unclear, we now ask: have all d/Egs had similar formation histories? Does mass to light ratio (T , in solar units) vary with luminosity?

3.4.1. Photometric properties. Consider first the azimuthally averaged radial SB profiles of Egs. While the azimuthal average may blur important differences between objects (e.g., the possibility that different kinds of Egs tend to show 'discy' versus 'boxy' isophotes), the radial SB profile is a natural starting point. Given the degree of uncertainty that arose from questioning the 3D structure of Egs in section 3, one should not be surprised that there is controversy over whether Egs form a single homologous family in their radial SB profiles $I(R)$. Much early work emphasized the difference between Eg and dwarf Eg profiles. Specifically, luminous (hence, massive) Egs are well fit by the de Vaucouleurs (1948) $r^{1/4}$ law

$$\ln(I(R)/I_e) = -7.67[(R/R_e)^{1/4} - 1] \quad (3.5)$$

with $I_e \equiv I(R_e)$ and factor 7.67 to ensure that half of the total intensity comes from $R < R_e$. On the other hand, dwarf Egs (and Sg discs) are fit better by the exponential profile

$$\ln(I(R)/I_e) = -R/R_e \quad (3.6)$$

A successful exponential fit does not imply that all Egs are actually disc dominated, although it does raise the question of whether some dwarf Egs masquerade as S0 galaxies. However, a different radial SB profile does indicate that d/Egs may differ fundamentally. Further evidence comes by plotting $\mu_0 \equiv I(0)$ the central SB against M_B , the logarithm of the blue luminosity of the galaxy. Starting from low luminosities, μ_0 increases monotonically with luminosity; at the higher luminosity end the trend reverses: μ_0 decreases as luminosity increases (for example Kormendy 1985). When few Egs with $M_B \sim -18$ had accurate surface photometry, it was generally thought that both the SB profiles and μ_0 's change character in a sharp transition near luminosity $M_B \sim -18$.

However, Caldwell (1983) showed that Egs form a continuous colour-magnitude sequence, without breaking at $M_B = -18$. In fact, recent papers (for example Trujillo *et al* 2001, Graham and Guzmán 2003) have advocated that Egs of all luminosities are fit by the more general Sérsic (1968) profile

$$\ln(I(R)/I_e) = -b_n[(R/R_e)^{1/n} - 1] \quad (3.7)$$

with b_n a function of the shape parameter n set so that half of the intensity is within $R < R_e$, and parameter n varies gradually from 1 at the low luminosity end to ≥ 4 (the de Vaucouleurs profile) at the high end. Graham and Guzmán (2003) find continuity in the plot of μ_0 (or μ_e) against M_B : Figure 5 shows that μ_0 first increases from low luminosity, then transitions smoothly near $M_B \sim -20.5$ to decrease with further increasing luminosity. They show that a Sérsic profile fits very well throughout the galaxy for all lower luminosity Egs; Graham (2005)

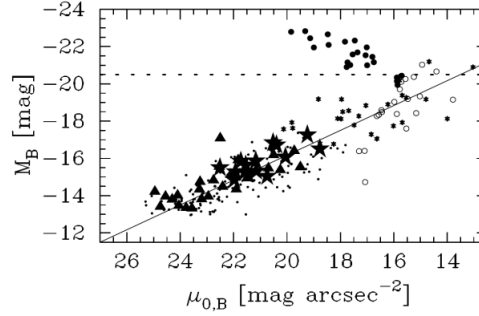


Figure 5. From Figure 2 of Graham (2005; used with permission), showing the reversing trend in central blue SB at the horizontal for Egs more luminous than log galaxy luminosity $M_B \sim -20.5$ in blue-magnitude units. The diagonal solid line (—) shows the correlation established from different data sets (different symbols, see paper) at low luminosity.

presents the case that Egs form a single family in their photometric profiles. Only at higher masses must it be modified in the inner regions to avoid over-predicting SB, and Graham and Guzmán (2003) argue that more massive Egs harbour a central supermassive black-hole, which Valluri and Merritt (1998) among others have shown may make orbits near the centre chaotic, thereby lower the central stellar mass density. However, Statler *et al* (2004) analyze one Eg studied by SAURON and show this does not appear to have happened.

What to make of these results? On the one hand, enough free parameters can always fit all galaxy profiles onto a single profile. On the other, the evidence of Graham and collaborators does alleviate the previously claimed discontinuity in Eg profiles from exponential to $r^{1/4}$, and also explains plausibly the transition in central SB as a function of galaxy luminosity. We therefore feel that, excluding dSph's, there is no longer a strong case to be made from photometric profiles that Egs divide into two distinct families.

3.4.2. Kinematics: the fundamental plane. Moving from photometry to kinematics, the tensor Virial Theorem has been applied to other properties of early-type galaxies that correlate tightly over at least R_e , if not globally. The space defined by SB I_e and velocity dispersion σ is the surface

$$R_e \propto \sigma^{1.24 \pm 0.07} / I_e^{0.82 \pm 0.02} \quad (3.8)$$

(for example Jørgensen *et al* 1996) called the ‘fundamental plane’ (FP). The FP shows that the final radius of an Eg inside its DM halo is set by the collapsing mass to within a factor of 2 over a hundredfold range of mass. The narrow range of possible final configurations of Egs is set by dissipation. A homologous collapse that ends in virial equilibrium with constant Υ would yield $R_e \propto \sigma^2 / I_e$. Systematics of the observed deviation from (3.8) through breakdown of homology — the so-called ‘tilt of the FP’ — show the dependence on galaxy mass or other global parameters of stellar Υ (hence the stellar population via the initial mass function, IMF) and stellar and DM phase space densities. If the dynamical basis of these systematics can be understood, then their trends with lookback time can provide insights into the mass build up of galaxies and can test models of galaxy formation and evolution.

How these dependencies vary among galaxies sets the intrinsic, thin scatter of the FP. For example, Jørgensen *et al* (1996) find that

$$\Upsilon_{\text{fp}} \propto \sigma^{0.86} \quad (3.9)$$

with 25% scatter. (Such small scatter yields fairly accurate distances, enabling study of peculiar velocities.) From extensive photometry by combining ground-based and HST I-band images of 25 galaxies and SAURON spectra within R_e , Cappellari *et al* (2006) find that Υ_{fp} can also be written within experimental uncertainties as a simple power law of either luminosity or mass alone. They derive mass densities from a Jeans model and distribution function that depends only on E and L_z , deprojecting their photometry and stellar velocities under an assumed axisymmetry and fixed Υ . They find only weak dependence of Υ on the assumed inclination angle i (good, because Krajnović *et al* (2005) show that i is constrained only weakly by IFS spectra). Next, fixing i from the Jeans model, they use Schwarzschild's orbit summing technique (Appendix A.2.4) to construct more realistic three-integral but still axisymmetric models that use more moments of the velocity function, not just the first and second used in the Jeans model. Fits to spectra yield the same constant Υ as from the Jeans model within scatter, demonstrating that the derived Υ is independent of the Jeans assumption of a homologous distribution function. The noisier Υ derived from their three-integral fits to the velocity function yields the tight correlation

$$\Upsilon(< R_e) = (3.8 \pm 0.14) \left(\frac{\sigma_e}{200 \text{ km s}^{-1}} \right)^{0.84 \pm 0.07} \quad (3.10)$$

Υ is found to depend mainly on galaxy mass and then, at fixed mass, on σ_e , a trend related to galaxy compactness. The scatter around (3.10) is small enough to exclude significant triaxiality for at least the fast rotators in their sample, supporting the assumptions that led to Υ . Examining the scatter as a function of luminosity, they find that the slow rotators, hence generally higher mass Egs of triaxial shape (section 3.3), have larger Υ than the fast rotators hence generally lower mass oblate Egs. They show that non-homology is responsible for $\lesssim 7\%$ of the variation in (3.10), most scatter coming instead from intrinsic variations of Υ inside R_e around typical stellar values that are consistent with simple stellar populations, see section 3.6.1) models. They find that the ‘virial mass’ defined as

$$M_{\text{virial}} = (5.0 \pm 0.1) R_e \frac{\sigma_e^2}{G} \quad (3.11)$$

is a reliable, unbiased estimator of mass within R_e . This result confirms that $R^{1/4}$ profiles and σ_e measured in a large aperture *can* measure mass build up even at high redshift.

3.5. Dark matter content

Cappellari *et al* (2006) compare their dynamical estimate of Υ with that inferred from simple stellar population synthesis to reproduce observed spectral indices. They reject a Salpeter (1955) IMF because the Kroupa (2001) form scatters less, and infer that the IMF barely varies among galaxies. They conclude that $\sim 30\%$ of the total mass within the R_e of early-type galaxies is dark, consistent with the few lensing cases (for example Ferreras *et al* 2005, Rusin and Kochanek 2005a). Only a variable DM fraction within R_e can explain the observed scatter.

Mateo (1998) reviews DM constraints from the internal kinematics of Local Group dSph galaxies. Masses and central velocity dispersions are usually derived using King models (Appendix A.2.2), which assume isotropic velocity dispersions and that mass follows light. Kleyna *et al* (2001) obtain radial velocities of many stars in the Draco dSph, find a flat/slowly rising velocity dispersion at large radii, hence argue for an extended DM halo. As a group, dSph's often have $\Upsilon \geq 200$, making them nearly as DM dominated as some low-SB galaxies (section 5.4.4). For example, Kleyna *et al* (2005) have studied 7 candidate stars in the faint, 100 kpc distant, recently discovered (Willman *et al* 2005) Ursa Major dSph. They identify five likely members, assume an isotropic velocity function with constant Υ , and obtain a

central $\Upsilon = 500$ solar from colours; they argue that this is a lower limit to the global value because the galaxy is so dark. Is this system a prototype of the ‘DM satellites’ that linger around mature galaxies at the end of Λ -CDM simulations (Navarro *et al* 1997)? Mateo (1998) obtained $\log \Upsilon = 2.5 + 10^7 / (L/L_\odot)$, i.e., that each early-type dwarf has a luminous component with $\Upsilon = 2.5$ solar that is embedded in a DM halo of $10^7 M_\odot$. Central black-holes more massive than $10^4 M_\odot$ are rejected in all systems except Ursa Major (Maccarone *et al* 2004) and Draco.

Further insights on DM have emerged from studies of planetary nebulae in Egs. A planetary nebula converts up to 15% of the white dwarf core’s radiation to the $\lambda 500$ nm [O III] emission line, allowing velocities to be measured accurately in early-type galaxies up to 30 Mpc distance on a 4-meter aperture telescope. They are important probes in more massive spheroidal galaxies because their emission lines are isolated easily from starlight with narrow-band filters, and their kinematics can be mapped very efficiently using a slitless spectrograph over a large field of view (Douglas *et al* 2002) and out to $5R_e$ to make application of the tensor Virial Theorem more reliable.

However, they can be < 3 Gyr old, a ‘young’ population in early-type galaxies. Thus their orbits have probably not relaxed dynamically to the same state as their parent galaxy, for example if their progenitor stars were formed in the late merger of a gas rich system. Romanowsky *et al* (2003) have studied ~ 100 planetary nebulae in each of four lower-luminosity Egs. They found that velocities decline by $\sim 60\%$ in Keplerian fashion between one and $3R_e$, obviating need for DM. However, they note that such a decline could also result from mostly radial stellar orbits, as might occur in an Eg formed when two discs merge. They derive $\Upsilon = 6.4 \pm 0.6$ solar that is consistent with their population synthesis, compared to 20–40 found in more luminous Egs. Napolitano *et al* (2005) obtain similar results. These observations seem to leave little room for DM within $2R_e$ in lower luminosity Egs.

However, challenged by such compelling data, Dekel *et al* (2005) choose more realistic stellar and DM density profiles, simulate disc galaxy mergers, find that they can reproduce the observed planetary nebula velocity function with $\beta(r > R_{\text{eff}}) \equiv (1 - \sigma_\theta^2/\sigma_r^2) \approx 0.75$ and a significant DM halo, and establish that such values arise from their simulations because radial orbits form preferentially in the outer parts of the merger remnant whatever the gas content of the progenitor galaxies. β correlates with tidal strength, being higher for more head-on collisions. It exceeds the ≤ 0.2 predicted for DM halos, suggesting that the planetary nebulae do not trace the DM. Merritt (1993) reached a similar conclusion, estimating that hundreds of nebular velocities would be needed to probe DM in the absence of other kinematical constraints. In fact, Pierce *et al* (2006) find that the velocity dispersion of globular clusters in NGC 3379 is instead constant with radius, arguing for a substantial DM halo. Clusters on initial radial orbits will soon disrupt tidally, reducing the orbital anisotropy of this population.

These studies emphasize the need for multiple DM tracers to provide consistency and highlight the uncertainties from dynamically ‘young’ populations. Using globular clusters and planetary nebulae separately is perilous.

How is DM distributed spatially? Equipotential surfaces are always more spherical than the matter distribution. It is therefore likely that a DM halo around a spheroidal galaxy is close to spherical. It may, however, be offset from the photometric centre of the galaxy, and in simulations is also found at small radii to be ‘cuspy’ (for example Navarro *et al* 1997). Kinematical evidence for a central cusp of DM is inconclusive because of uncertain stellar disc and bulge mass decompositions, the presence of distinct dynamical subsystems in the core, and plausible variations in Υ , dust and stellar populations throughout the volume.

3.6. Stellar Content

Having discussed the structure and dynamics of Egs, we turn to questions that bear on their formation and evolution.|| Unfortunately, despite growing capabilities to detect and study the luminous content of galaxies to lookback times of their formation, we cannot observe the assembly of the dominant DM halos. Instead, we must rely on observations of the SF and chemical evolution histories of the luminous baryons to infer how underlying DM evolves. Given compelling evidence from both observations (for example Searle *et al* 1973, Larson *et al* 1980) and numerical models (for example Mihos and Hernquist 1996) of galaxies undergoing gravitational interactions/mergers, it is generally agreed that an Eg that forms by merging anything other than two gas-free progenitors will show enhanced SF and chemical evolution. So, a key question is: to what extent can SF/chemical evolution histories of galaxies be extracted from their spatially integrated starlight? In particular, how well can one determine the mean age and chemical composition of stars in an Eg? And, if SF indeed continued after a few dynamical times as the DM halo formed, how reliably can one measure the SF/chemical evolution history? Naturally, to measure discontinuities in this history would be particularly useful.

3.6.1. Key issues for integrated light spectroscopy. Before addressing results from integrated colours and spectra, we highlight some important complications.

- The evolution of a coeval stellar population is highly non-linear in time. That is, the integrated spectrum first evolves very rapidly, then slows as the MS turn-off (see Appendix A.1.5) drops to lower masses. In fact, the models in Figure 6a show that the population ages logarithmically.
- The integrated Υ of the evolving population also increases dramatically and logarithmically from very small values for a young population. Thus, even limited recent SF tends to overwhelm light from the old population, skewing sharply the luminosity-weighted mean age of a galaxy toward the age of the youngest stars. Such bias is, of course, useful if one wishes to detect even small amounts of recent SF from mergers.
- Degeneracy between age and metallicity was the most serious complication to confront the first studies in integrated light. Specifically, Figure A1 shows that an older, metal-poorer population mimics a younger metal-rich one. However, as detailed below, measuring mainly age-sensitive hydrogen Balmer absorption lines coupled with metal features such as Fe and/or Mg lifts this degeneracy (Worthey 1994). Age/metallicity and dust reddening are a second degeneracy that alters photometric colours and the overall spectral energy distribution. Dust reddens starlight, making a galaxy look older and/or metal-rich. However, individual spectral features are unaffected.
- The most vexing hence interesting current issue for integrated light studies is that element abundance ratios in Egs do not track those in the Sun and Solar Neighbourhood stars. As we discuss below, recent work shows that the α -element/Fe ratio exceeds solar in Egs, indicating more rapid chemical evolution than occurred locally. While this result provides rich insights on galaxy evolution, it unfortunately complicates models of integrated spectra.
- Extracting a galaxy age from an integrated spectrum demands far more of stellar evolution models than extracting the age of a resolved star cluster from its CMD. For a

|| A reader unfamiliar with the basics of stellar evolution should first review Appendix A.1 where some standard abbreviations are defined.

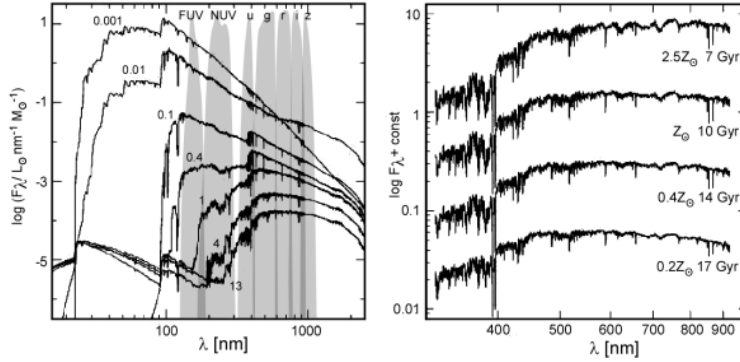


Figure 6. (a) Evolution of the standard simple-population synthesis model of (Bruzual and Charlot 2003; used with permission) for solar metallicity, using their GALAXEV models (<http://www2.iap.fr/users/charlot/bc2003/index.html>). Times since the end of the star burst are given in Gyr; note very small changes 4 – 13 Gyr after the burst. Between 0.1 – 1 Gyr, all Balmer absorption lines strengthen from late-B to early-F stars; this is the standard diagnostic of a recent star burst. Over the same interval, the ‘Balmer break’ from the Balmer continuum limit evolves into the ‘4000 Å break’ from numerous prominent metal lines to the blue. In gray are the *GALEX* FUV, *GALEX* NUV, and SDSS filter bandpasses. Sometimes, ages are determined crudely from the relative intensities of light in adjacent bandpasses. (b) A series of simple-population synthesis spectral models of increasing age and decreasing metallicity (top to bottom) show how the two parameters are nearly degenerate. However, close examination of the spectra reveals that the apparent degeneracy can indeed be broken with spectra that resolve individual line strengths.

CMD age, stellar modellers match only the position and shape of the MSTO (Appendix A1.5). In contrast, an integrated spectrum is a composite of light from many parts of the HR diagram. So, the shape of isochrones must be correct across the HR diagram, not just at the MSTO. Even more challenging, isochrones must have the correct stellar *number density* across the HR diagram. Equivalently, the timescales for *all* advanced phases of stellar evolution must be correct. For example, Schiavon *et al* (2002a;b) show that discrepancies between the age of the MWg globular cluster 47 Tuc derived from its integrated light and from its CMD may arise largely because isochrones under-predict stars along the upper part of the RGB.

3.6.2. Results for the dwarf elliptical galaxy M32. Given these complications, the reader should question whether reliable mean ages and chemical abundances of Egs can be extracted from integrated spectra and photometric colours. In fact, this issue has been controversial for decades. Pioneering models (Spinrad and Taylor 1971, Faber 1972) discovered that Eg integrated spectra are dominated by the light of metal-rich, mostly old stars. Moreover, there were strong indications that successful models require stars with CN and Mg features that are only reproduced by stars in the Solar Neighbourhood whose line strengths exceed greatly those found in stars of solar abundance, a result that is now seen in the context of enhanced-to-solar α -element/Fe abundances. Faber (1972) found that the mean metallicity of Egs increases with mass. However, these studies could not constrain Eg ages beyond showing that Eg spectra are clearly dominated by older stars.

O’Connell (1980), in another landmark study of galaxy spectra, analysed the integrated spectrum of the nucleus of the dwarf Eg M32 and concluded that the Balmer absorption lines are too strong to be explained just by an old, metal-rich population. Rather, M32 must contain

MSTO stars of intermediate age. Clearly, that Egs can contain such stars has important implications on their formation and evolution, so O’Connell’s controversial assertion has been much scrutinized and has made M32 a ‘Rosetta Stone’ for integrated light studies of galaxies. In fact, its nucleus is bright enough to obtain a very high SNR spectrum ($\text{SNR} > 100$ per $\lambda 0.1$ nm) in only a one-hour integration on a small, 1.5-meter aperture telescope. Moreover, due to its low ($\sim 80 \text{ km s}^{-1}$) central velocity dispersion, spectral features are better resolved than in the centres of massive Egs (central velocity dispersion $\sim 300 \text{ km s}^{-1}$).

Several ways have been proposed to enhance metal and Balmer absorption lines without resorting to a multi-age population (see Renzini 1986; for a summary). Specifically, populations with a range in metal abundance can reproduce the observed spectrum, i.e., an age-metallicity degeneracy argument. Alternatively, a few hot ($\sim 10,000 \text{ K}$) stars could strengthen Balmer absorption lines without resorting to a major intermediate age population; these might be blue stragglers — upper MS stars from a very minor ~ 1 Gyr old population — or blue HB stars from a metal-poor population. Finally, perhaps non-solar abundance ratios have so contaminated the bandpasses of the indices used to measure key spectral features that we are not modelling properly the integrated spectrum.

Clarifying the situation of M32 and other Egs led to key advances. Tinsley and collaborators (for example Gunn *et al* 1981) introduced the now standard evolutionary population synthesis, i.e., comparing integrated spectra of galaxies to a grid of model populations based on isochrones and either empirical or synthetic spectra of stars. (Synthesis replaced optimized fits to a galaxy spectrum from an observed spectral library with few stellar astrophysics constraints.) Another advance was the development of the Lick system of spectral indices (for example Faber *et al* 1985), a set of passbands defined to isolate and quantify important spectral features. In a key advance, Worthey (1994) modelled carefully the age-metallicity degeneracy. He showed that plotting a Balmer line index (e.g., strength of $\text{H}\beta$) against a metal-sensitive index (e.g., strength of the $\text{Fe I } \lambda 5270$ feature) decouples the age-metallicity degeneracy. Finally, population synthesis models have begun to incorporate non-solar abundance ratios (NSAR) to model spectra and indices (Maraston *et al* 2003). This requires modelling the effects of NSAR on the absorption features due to the abundance changes on individual features (Tripicco and Bell 1995), and on the internal stellar structure and evolution due to changes in, for example, the interior opacity from changing proportions of major electron donors such C, N, and O.

The present situation is encouraging. Simple stellar population models now produce sensible ages and metallicities for well understood stellar systems such as globular (Schiavon *et al* 2002a;b) and open clusters (Schiavon *et al* 2004). For M32, recent analyses agree that an intermediate age population is unavoidable to reproduce the integrated spectrum of its central ~ 3 arc-seconds, both Worthey (2004) and Rose *et al* (2005) finding that the light-weighted mean age doubles from the central value ($\sim 3\text{--}4$ Gyr) at the nucleus to R_c . Mean metallicity halves, so these competing effects ‘conspire’ to keep the integrated B-V colour roughly constant. Rose *et al* (2005) demonstrate with a spectral indicator sensitive to light from hot stars that the enhanced Balmer absorption lines (over that expected for a uniformly old population) cannot be attributed to a small population of hot stars. Hence, a more substantial intermediate age population is indeed necessary to explain the strong Balmer lines. Moreover, both Worthey (2004) and Rose *et al* (2005) find only modest departures from solar abundance ratios in M32, simplifying the interpretation of its integrated spectrum.

3.6.3. Results for large samples of elliptical galaxies. We turn to results from large surveys of spatially integrated Eg spectra. Trager *et al* (2000a) studied 50 Egs in both the low-density field and clusters, and found a significant range of light-weighted ages. The range is widest

for field Egs and for those with lower velocity dispersions σ . As well as being older on average, larger σ Egs have larger NSAR, with [Mg/Fe] enhanced heavily in the most massive ones. Caldwell *et al* (2003) obtain similar results in a study of nearly 200 Egs that cover a wide range in σ (hence luminosity and mass). Most striking is their result that low σ galaxies scatter in age far more than those with high σ , i.e., low-mass Egs have had more prolonged SF histories. In contrast, in a study of early-type galaxies in the Fornax cluster, Kuntschner (2000) found that all Egs have similar old ages, but that metallicity increases with σ . In a recent study, Sanchez-Blazquez *et al* (2006) analyse integrated spectra of nearly 100 field and cluster Egs. They find that Egs in low-density environments range in age more than their clustered counterparts, with those in low-density environments tending to be more metal rich and younger by ~ 1.5 Gyr. For these galaxies, an age-metallicity relation is found when metallicity is measured by Fe (but not by Mg). While a well-defined mass-metallicity relation is found for clustered galaxies from both Fe and Mg, for field galaxies a mass-metallicity relation is only evident in Mg. They also find that lower luminosity galaxies spread more in age than their higher luminosity counterparts.

Remarkably, these studies have uncovered little evidence for a mass-metallicity relation among field Egs. Most lower mass Egs have had prolonged SF histories and slower chemical evolution than their higher mass counterparts. In contrast, clustered Egs are systematically older than Egs in the field, hence much more consistent with the hypothesis that SF occurred rapidly and ended early to produce a well-defined mass-metallicity relationship.

How do these results relate to predictions of the Λ -CDM hierarchy? As mentioned, one cannot simply compare galaxy SF/chemical evolution histories to the simulated build up of CDM halos. Rather, simulations must incorporate parametrized simplifications of baryonic astrophysics to follow luminous particles. For example, Somerville and Primack (1999) have predicted SF histories of Egs, finding, unsurprisingly, that they follow the assembly timescale of DM halos. Specifically, at all masses (luminosities) one expects from the stochastic build-up of the hierarchy to produce a range in light-weighted mean ages. Too, more massive galaxies tend to have prolonged SF (hence younger light-weighted mean ages) than low mass galaxies. Moreover, at given mass, SF is prolonged for galaxies in low density field environments compared to in rich clusters, as predicted by Kauffman and Charlot (1998). As Proctor and Samsom (2002) and Sanchez-Blazquez *et al* (2006) discuss, it is encouraging news for Λ -CDM models that light-weighted mean ages of Egs spread significantly, and also that Egs in clusters are, in the mean, older than those elsewhere. However, that mean ages of lower mass Egs are younger (and less enhanced α -elements) than for massive ones counters a straightforward prediction of hierarchical models.

Inverted observation/model trends clearly require a different model prescription for SF feedback, i.e., something that quenches SF in the deeper potentials of more massive galaxies. It has been argued recently (for example Benson *et al* 2003, Bower *et al* 2006) that the solution to this problem and to others in predicting the form of the galaxy luminosity function at both low and high redshift, is connected with the formation of supermassive black-holes in the centres of more massive galaxies and the subsequent quenching of SF (see section 6).

Finally, one must recognize that existing data cover only the centres of Egs. Light-weighted mean ages and metallicities are dominated by the nuclear region in the spectrograph slit or few arc-second diameter fibre, generally only a small fraction of R_e . Long-slit studies of M32 indicate that the intermediate age population concentrates centrally (Worthey 2004, Rose *et al* 2005). Dwarf Egs in the Virgo cluster, which typically have a younger light-weighted mean age (Caldwell *et al* 2003), are bluer in the nucleus (Vader *et al* 1988), an ‘inverse’ radial colour gradient. Together, these results argue that the young and intermediate age populations in Egs may be mostly nuclear. Unfortunately, little is known of population

gradients in Egs, due partly to centrally concentrated light that makes extremely challenging the high SNR studies out to R_e necessary to disentangle age from metallicity. Radial colour gradients traced to faint SB (for example Vader *et al* 1988, Peletier *et al* 1990, Wu *et al* 2005) reveal that massive Egs have redder centres. The straightforward interpretation is that their nuclei are more metal-rich. However, disentangling age from metallicity with only visible-band colours is problematic. With rapid progress in near-IR imaging, adding near-IR colours increases discrimination between age and metallicity thirtyfold (Cardiel *et al* 2003, MacArthur, Courteau, Bell and Holtzman 2004). (We caution that modelling near-IR spectra requires understanding all the issues of stellar mass loss and convection discussed in Appendix A.1.5; see Lee *et al* 2007.) Kobayashi and Arimoto (1999) summarized work on spectral index gradients in Egs from many sources. For massive Egs at least, they concluded that universal NSAR's (as characterized by $[Mg/Fe]$ and significant metallicity gradients) exist.

3.6.4. Stellar populations: the next steps. The bottom line from the previous section is that current data support hierarchical formation in some ways, and challenge it in others. What near-term, perhaps decisive, improvements in both observations and modelling of integrated spectra can we anticipate? There are two data improvements. First, results to date have solved the first-order problem of the light-weighted mean ages and chemical compositions of Egs. Now we must discriminate between multiple populations. Specifically, to what extent is the 3 – 4 Gyr mean age of M32's nucleus a composite old and young population, or perhaps instead multi-episodic? Second, as mentioned above, data on radial population gradients to constrain scenarios of Eg assembly are sparse, especially for low mass Egs.

The first question of distinguishing mean ages and metallicities in one or more populations is addressed by using multiple spectral features that range widely in wavelength. A younger population, being bluer than the older, will contribute proportionately more blue light than red. Thus light-weighted mean ages from bluer spectral indices give younger ages than those from redder ones (for example Sanchez-Blazquez *et al* 2006, Schiavon *et al* 2004). Such subtle distinctions require very high SNR and very reliable models of different spectral indices. For example, clear problems in modelling the $H\delta$ line as an age indicator come from 'contamination' by a nearby CN molecular band that can only be modelled confidently by using the correct NSAR prescription (Prochaska *et al* 2007). So the problem is not strictly one of better quality data; we also need better models.

Much modelling effort is underway to construct fully consistent NSAR grids that incorporate not only the stellar interior/evolution effects of NSAR, but also effects on the atmospheric structure and its emergent spectrum. Ongoing work to synthesize spectra for various NSAR prescriptions will eventually replace empirical spectral libraries and their limitation to Solar Neighbourhood abundance ratios (Coelho *et al* 2005).

Two tactics yield reliable radial age and abundance gradients in galaxies. First, measure gradients out to interesting radii by going as faint as is possible reliably. This requires not only many photons hence large-aperture telescopes and high-throughput spectrographs and detectors, but also very accurate sky subtraction and control of scattered light. Second, cover an extended region with an IFS, do not view a narrow slice of it through a slit.

The second advance is the advent of vast spectral surveys such as the 2dF and the SDSS. Recent work on extracting the SF histories of many galaxies in the SDSS have not targeted Egs (Jimenez *et al* 2005) specifically. In fact, fingering Egs for separate analysis is unwise because morphology is ephemeral in the Λ -CDM hierarchy. As discussed above, with exceptional SNR on a single galaxy across many different radii, one might distinguish multiple episodes of SF from a light-weighted mean age. That would be a true advance. A clever example of constraining SF by statistics is Trager *et al* (2000a), who posit a 'frosting' model for

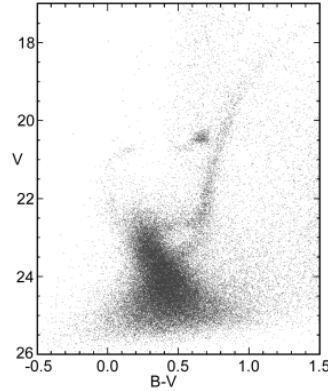


Figure 7. CMD of the centre of the Carina dSph, a MWg satellite, showing multiple sub-giant branches, and blue and red HB stars (Monelli *et al* 2003; used with permission). The vertical axis is related to log visible-light luminosity, and the horizontal axis is colour with blue at left and red at right. Isochrone fits yield distinct populations aged 0.5, 5, and 11 Gyr from the present if helium and metal abundances are $Y = 0.23$ and $Z = 0.0004$, respectively.

Egs to explain the distribution of light-weighted mean ages in their sample. Frosting adds one episode of young stars to an old population. Leverage increases when galaxy ages and metallicities from intermediate and high redshift surveys (Faber *et al* 2005; for example DEEP-2 and COMBO-17) are combined with current epoch large samples like the SDSS. Clearly, constraints strengthen when statistical properties of galaxies must be reproduced at multiple epochs. It is beyond our scope to review the growing observations of galaxy evolution at high redshift; we mention only that recent data prefer ‘quenched’ SF (SF that ends abruptly rather than a single recent burst) over frosting (Faber *et al* 2005).

3.6.5. The lowest luminosity (Local Group dwarf) spheroidal galaxies: resolved stellar populations. In testing models to form and evolve Egs, much is gained by studying the smallest and faintest galaxies because they are the most common. Hundreds of dwarf Egs have been catalogued in the nearby Virgo (Binggelli *et al* 1985) and Fornax (Ferguson 1989) clusters. To test the Eg model in the extreme, one can study individual stars in Local Group dSph’s. dSph’s are very low luminosity ($M_V > -14$), low central SB ($\mu_{0,V} > 22$ mag arcsec²), gas-poor galaxies that barely rotate. Their relation to more massive Egs and to dwarf Egs (which are somewhat more luminous and have brighter centres than dSph’s) is unclear. Mateo (1998) and Grebel (2005) review their properties. In galaxy groups, most lie within ~ 300 kpc of a massive galaxy, whereas dIrr’s are more dispersed (Grebel 2005).

SF and chemical evolution histories of Local Group dSph’s can be detailed with ground-based and HST multi-colour photometry to provide CMDs, and by using large aperture telescopes to obtain intermediate-resolution spectra of ~ 100 stars simultaneously (for example Venn *et al* 2004). With metallicity from spectra, one can then fit isochrones to broad-band photometry to pin down the SF history. Of course, a complex SF history will superimpose multiple isochrones, increasing the likelihood of a non-unique decomposition (Gallart *et al* 1999), for example Figure 7.

Key results come from CMD and metallicity studies. First, except for the Ursa Minor dSph (Carrera *et al* 2002), all Local Group dSph’s show intermediate age stars hence prolonged SF histories (Koch *et al* 2007, Dolphin 2002, Saviane *et al* 2000, Gallart *et al* 1999). This is very surprising because energy from the first ~ 1000 SNe in the first star

burst exceeds greatly the binding energy in the shallow potential of the galaxy (total mass only $10^7 M_{\odot}$), implying the ejection of all ISM then. Consider the Carina dSph, the only such system whose SF can be resolved: Figure 7 shows that it formed stars < 1 , $3 - 6$ ($\sim 70\%$ of total), and $11 - 13$ Gyr ago with few between. Recent VLT spectra yield mean metallicity $[Fe/H] = -1.7$, ranging from -3.0 to 0 . Metal-rich stars concentrate toward the centre. Carina has $Y \sim 32$ with uncertain DM content and essentially no gas. Multiple star bursts mean either that gas ejection was inefficient or that new gas was accreted. Accretion or re-accretion would be difficult because of this dSph's very low total mass, small cross section, and high orbital speed in the MWg's outer halo ($\sim 200 \text{ km s}^{-1}$). So, merging gas would have had to move in nearly the same orbit. Or, gas ejection may not remove the densest giant molecular clouds (GMC) where SF occurs. What causes the few Gyr latency between bursts is unknown.

A second indicator of a history of chemical evolution is the discovery that dSph's like Carina range substantially in chemical composition. As expected, mean metallicity increases with luminosity in dSph's. Grebel, Gallagher and Harbeck (2003) among others have noted that the metallicity-luminosity relation in dSph's offsets to higher metallicity at a given luminosity compared to dIrr's (gas-rich low-luminosity galaxies, generally with substantial rotation), thereby questioning if the latter can be precursors of dSph's.

Third, is the question of whether the MWg halo, particularly its component in retrograde motion, could have assembled from accreting dSph's. Venn *et al* (2004) have analysed detailed abundance patterns in the MWg halo and in dSph's. They find that dSph's have lower $[\alpha/Fe]$ at given $[Fe/H]$ than halo stars. There is better consistency with halo stars in extreme retrograde orbits, but other element abundance ratios do not agree. Higher $[Ba/Y]$ in dSph stars compared to that in halo stars indicates slower chemical evolution in dSph's, because Ba forms in the s-process. This result accords with the extended SF histories of dSph's. So, while the halo is clearly not assembled from stars with the chemical evolution of today's dSph's, it is unclear if earlier mergers are excluded entirely, especially because the most metal-poor stars in dSph's appear to have abundance patterns more consistent with the halo.

Finally, the distribution of Local Group satellites is interesting. Hartwick (2002) finds evidence for dissipational collapse from the anisotropic velocity ellipsoid of this population: it is flattened spatially along the major axis, indicating a non-spherical potential. M31 and its satellite retinue including M33 show a similar alignment (Koch and Grebel 2006). Hartwick argues that such is expected when a flattened matter filament collapses rapidly to a plane as the globular cluster and dwarf galaxy population formed. The collapse then progresses in the plane to form the giant galaxy (Larson 1969). Azzaro *et al* (2007) examine the alignment of satellite galaxies in the SDSS dataset, and find that red ones are aligned along the major axis of red hosts up to 0.5 Mpc away; the sample is too small to test for isotropy around blue hosts.

Thus there is no consensus yet on the origin of dSph's because of their surprisingly large spread in stellar age and chemical composition. The nature of these extremely low mass galaxies remains a particularly important challenge for any scenario of galaxy evolution.

3.6.6. Local Group globular clusters. The next rung down in mass and our main probe of baryonic halos beyond the Local Group are globular clusters. There are ~ 500 in M31, and ~ 150 in the MWg ¶ with continuing discoveries from near-IR surveys. Some Egs have more than 10,000. HST or an adaptive optics system resolve all but the most distant Local Group globular clusters to their core. Not all are thought to be bound in galaxies, some may be 'intergalactic tramps'. While there are counter examples of globular clusters younger than 1 Gyr in the LMC and especially in merging systems, the MWg, M31, LMC, Fornax dSph,

¶ See the tabulation at <http://physun.physics.mcmaster.ca/Globular.html>

and Sagittarius dIrr's all have globular clusters of comparable ancient age within ± 1 Gyr. An exception is the massive cluster ω Cen, which has multiple stellar populations and anomalous chemical abundances (especially copper) compared to other globular clusters.

Globular clusters in most galaxies have a bi-modal colour distribution (for example Ashman and Zepf 1992), usually interpreted as two populations and the signature of a hierarchical build-up. However, Yoon, Yi and Lee (2006) have challenged this interpretation by noting the strong non-linear dependence of globular cluster colour on metallicity through the dependence of HB morphology on metallicity. They argue that the colour distribution is due instead to a continuous range of metallicity in a single old population, thereby avoiding multiple metallicity and/or age populations in globular cluster distributions.

3.6.7. Local Group archaeology: chemo-dynamical clues in Local Group baryon halos.

Local Group archaeology aims to sequence the events that formed those galaxy discs and halos, and thereby to constrain models of galaxy formation. Including recently identified objects, the MWg baryon halo has $\sim 3 \times 10^9 M_{\odot}$ within 50 kpc radius ($< 0.6\%$ of the total), density profile $\rho(r) \propto r^{-3.5}$ out to radius ~ 100 kpc, and may have two components (inner Halo I moderately flattened, outer Halo II more spherical according to Gould *et al* 1998). Analyses of both globular clusters (Rodgers and Paltoglou 1984, Lee and Carney 1999) and halo field RR Lyraes (Borkova and Marsakov 2003, Kinman *et al* 2007) and subdwarfs (Majewski *et al* 1996) identify two kinematic components to the halo: one in mean retrograde motion, the other in slow prograde rotation (Chiba and Beers 2001, Reid 2005). The former is believed to result from tidal disruption of accreted satellites. The key question of whether most of the halo formed in a single monolithic collapse, (Eggen *et al* 1962), or through the more stochastic accretion of fragments (Searle and Zinn 1978), is still under debate (for example Chiba and Beers 2001, Gratton *et al* 2003).

Phase space structure in the halo may be preserved, but the relative importance of dissipation over accretion during halo formation is unknown. With enough stellar kinematical data, Freeman and Bland-Hawthorn (2002) argue that one can identify the stream from each accreted satellite. Few stars have been observed at sufficient spectral resolution for this, and in any event spectra can today be obtained only of inner halo objects. In the Solar Neighbourhood only 0.5% of the stars are in this component, a few thousand high-velocity or metal-poor stars (Carney *et al* 1996). Halo I and II stars are 10 – 13 Gyr old, but the bulge seems to be younger, 8 – 11 Gyr. There is a large range in $[\text{Fe}/\text{H}] = -5$ to $+0.5$, but element abundances in Halo I differ clearly from those measured in Local Group dSph's.

Photometry is much more extensive, for example from the SDSS. At $V > 18$, density enhancements in Halo II are large (Ibata *et al* 2003, Newberg *et al* 2002), especially tidal debris from the Magellanic Stream plus Sagittarius dwarf galaxy (our nearest neighbour galaxy, Majewski *et al* 2003a, Majewski 2004, Belokurov *et al* 2006) that together span the sky. A few moving groups have been identified ($1^{\circ} \times 10^{\circ}$ stream associated with globular cluster Pal 5, Odenkirchen *et al* 2003), and the mean and dispersions of their metallicities and kinematics have been measured; the Monoceros stream has been interpreted either as an accreted galaxy at ~ 15 kpc Galactic radius (for example Helmi *et al* 2003, Newberg *et al* 2002) or as part of the MWg's warped disc (Momany *et al* 2004).

M31 is the other Local Group galaxy with a detectable halo (the halo of M33 is elusive, McConnachie *et al* 2006, but see Ibata *et al* 2007). The V – and i -band Isaac Newton Wide-Field Camera Survey of M31 resolves stars 16 times less luminous than those at the tip of the RGB and extends out to 55 kpc radii (Ibata *et al* 2001b). The Canada-France-Hawaii Telescope MegaCam survey goes deeper and covers a larger field. In these images Ibata *et al* (2001b) and Ibata *et al* (2007) find a half-dozen star streams in that bulge/halo, including

one at 70 kpc (the Giant Stream, comparable to the Magellanic Stream) that extends for 100 kpc and whose large velocity deviation indicates no association with either of M31's large satellites M32 and NGC 205 (Chapman *et al* 2006). Assuming an orbit apocentre of 125 kpc, they find (Ibata *et al* 2004) $7.5^{+2.5}_{-1.3} \times 10^{11} M_{\odot}$ halo mass within. Guhathakurta *et al* (2006) study the halo at 10 km s⁻¹ resolution and obtain metallicities from the Ca II triplet. Chapman *et al* (2006) detect a distinct, non-rotating, metal poor ($[Fe/H] \sim -1.4 \pm 0.2$) stellar component from 10 to 70 kpc radius, of abundance similar to Halo I of the MWg (Chiba and Beers 2001). Evidently, M31 has had a more exciting life than the MWg, including a surprisingly recent merger (0.25 Gyr according to simulations by Font *et al* 2006).

In summary, the seeming simplicity of Egs once suggested that they are homologous in mass, largely supported by random motions but flattened by rotation, and assembled quickly (a few collapse times) while forming stars efficiently. Improved observational tools allowed study of 'ellipsoids' of all masses, and we have seen in this section that none of these assumptions have turned out to be true.

4. Bulges, pseudo bulges and bars in disc galaxies

An Sg generally contains a spheroidal stellar component that is more centrally concentrated than its flattened disc. This 'bulge' provides clues for galaxy evolution. As section 3 explained, the Λ -CDM scenario assembles a bulge through a near-equal mass merger. The result is an Eg until it reacquires a disc by cooling the residual diffuse halo of ISM that was heated by the merger. Hence in the Λ -CDM scenario, Sg bulges should resemble Egs in their internal dynamics and stellar populations. So, is a massive bulge also flattened mainly by an anisotropic velocity dispersion tensor? What has the bulge of the MWg told us about other bulges? Are bulge properties consistent with hierarchical mergers, or are other formation processes relevant?

To answer, one must address two issues that regard the bulge as a separate entity from the often dusty disc. First, we must separate their light profiles. Large bulges show the centrally concentrated SB profiles that characterize Egs, a Sérsic profile with index $n \sim 4$. Disk light declines in an exponential

$$I(R) = I_0 \exp(-R/R_d) \quad (4.1)$$

of I_0 central SB and R_d scale length. Sometimes the disc dominates at large radii so can constrain R_d while the bulge dominates at small. When separating disc and bulge light, the outer fit is then extrapolated inward to constrain I_0 and is subtracted from the observed profile; the residual is fit with a Sérsic profile to define the bulge (for example Schombert and Bothun 1987, MacArthur *et al* 2003). Thus, the bulge profile is only reliable for a pure exponential disc profile. 2D surface photometry is required and, in early-type (i.e., bulge-dominant) Sgs, projection effects must be considered (Noordermeer 2006). Dust grains attenuate bulge starlight. In the B-band, Driver *et al* (2007) find that only 29% of the bulge starlight escapes into the intergalactic medium.

4.1. Pseudo bulges

Second, it seems that not all bulges arise from mergers. Instead, as reviewed by Kormendy and Kennicutt (2004), 'pseudo bulges' evident in disc-dominant (so-called late-type) Sgs are often built by secular processes; Noordermeer (2006) finds similar flattened bulges even in some early-type Sgs. Specifically, non-axisymmetric disturbances (bars) can trigger gas inflow and subsequent star formation to build up a pseudo-bulge. That its scale length is well-correlated

with that of the disc (Courteau *et al* 1996, MacArthur *et al* 2003) further suggests its secular origin.

Other structures that at low spatial resolution mimic ‘classical’ bulges (but resolve into more complex structures in HST and adaptive optics system images) are inner discs, bars within bars, and box-shaped nuclear bulges (perhaps nuclear bars seen edge-on). Near-IR surveys that are sensitive to the stellar mass distribution have found many bars in visible-light unbarred Sgs. For example, the bulge of M31 is barred (Beaton *et al* 2007), and is twice as massive as that in the MWg, perhaps (Brown *et al* 2006) having been augmented by a merger with a galaxy of MWg mass. Kormendy and Kennicutt (2004) conclude that there are comparable numbers of early- and late-type galaxies, classical bulges and pseudo bulges. However, pseudo bulges in late-type Sgs have masses only 1 – 10% of bulges in Egs. Driver *et al* (2007) find from their Millenium Galaxy Catalog that 60% of stellar baryons lie in galaxy discs, and 27% in classical bulges.

Carollo (2005) reviews photo-kinematical distinctions between classical and pseudo bulges, finds that they are quite blurred, and challenges their division into distinct components. To the extent that this distinction is tenable, the consensus is that the central light excess over an inward-extrapolated exponential disc tends, in late-type Sgs, to be a pseudo bulge from secular evolution, but is a classical bulge in bulge dominant Sgs.

4.2. Bars and other non-axisymmetric distortions

As was touched on above, bars and other non-axisymmetric distortions can play an important rôle in the secular evolution of pseudo bulges. They are reviewed by Sellwood and Wilkinson (1993), Buta and Combes (1996; focused primarily on galactic rings, but discussing bar dynamics too), Knapen (1999), Knapen *et al* (2000), Shlosman (2001), and Kormendy and Kennicutt (2004). Specifically, bars drive gas inward to form stars, and they scatter disc stars into the pseudo bulge. Here we address only questions related to the issue of bulge versus disc development: How are bars generated? How long do they last? How often are they found?

Bars form either through internal secular processes or through external tidal triggering. Under appropriate conditions set by the mass distribution in the disc and halo, and the disc velocity dispersion, gas discs are unstable to bar and spiral structure formation through the ‘swing amplification’ mechanism (for example Toomre 1981). In brief, low-amplitude leading spiral wave patterns propagate through the galaxy centre and emerge as amplified trailing waves. A bar develops by exchanging angular momentum with the outer stellar disc, DM halo, and/or gas disc. In hot stellar systems (e.g., Egs and stellar interiors), any energy exchange produces a bifurcated core-halo structure (Lynden-Bell and Wood 1968) because of the negative specific heat of a self-gravitating system. Similarly, its angular momentum components separate (Lynden-Bell and Kalnajs 1972, Kormendy and Fisher 2005), with resonance between the rotation frequency and the frequency of radial oscillation delineating the in/outward flows of angular momentum. Extensive explanations for bar formation based strictly on orbit theory are provided in, for example Sellwood and Wilkinson (1993), Buta and Combes (1996). Naturally, numerical simulations of spiral discs are a crucial reality check on simplified Hamiltonians. Early simulations were restricted to 2-D, and were sufficiently grainy due to the small number of particles tracked that the secular bars that formed could plausibly have been artifacts of numerical instability. However, with greatly increased sophisticated numerical models now in 3-D, it is clear that real bar instabilities occur and that their strength depends on the mass distribution and velocity dispersion of the disc and halo (for example Knapen 1999, Shlosman 2001).

While bar instabilities may form in isolated discs only through internal processes, tidal

interactions and minor mergers are also effective triggers to drain angular momentum from the gas and inflow it to perhaps form stars and build up a pseudo bulge. Thus it is of considerable interest whether most pseudo bulges have been created by external (interaction/merger) processes, which would be more consistent with the hierarchical merger picture (Kannappan *et al* 2004), or by internal processes, which would pose problems for that scenario (Kormendy and Fisher 2005). In support of the external trigger, Kannappan *et al* (2004) find a correlation between central blue colours in pseudo bulges with morphological evidence of tidal encounters or merging. On the other hand, Kormendy *et al* (2006) examined two pseudo bulges in detail, and found no evidence for an external trigger.

The longevity of a bar depends upon the importance of gas in its evolution (for example, the simulations of Debattista *et al* 2006). For purely stellar systems, bars tend to be long-lived, in fact are only weakened by tides (Athanasoula and Bosma 2003). This is because angular momentum flows from the bar to either the outer disc or the DM halo, which slows the pattern speed and increases the strength of the bar (for example Athanasoula 2003, Athanasoula and Bosma 2003). On the other hand, when gas is included, there is general consensus that its interaction with the bar weakens the bar, but little agreement on how this occurs. Bournaud *et al* (2005) propose that the bar drains angular momentum from the gas, which increases bar pattern speed and weakens the bar. Berentzen *et al* (2007) argue that the gas has little direct effect. Instead, the bar is destroyed by the build up of the pseudo bulge as gas inflows to form stars, a less direct rôle for gas than Bournaud *et al* (2005) propose. The timescale for bar destruction is similarly unclear, with some investigators finding a quick $\sim 1\text{--}2$ Gyr one (Friedli *et al* 1994, Bournaud *et al* 2005), and others much longer (for example Berentzen *et al* 2007).

Finally, observations in the near-IR H band, which better reveal the bulk of the underlying stellar distribution than do optical passbands, find that $\sim 2/3$ of Sgs show a bar (Knapen 1999, Knapen *et al* 2000, Eskridge *et al* 2000, Whyte *et al* 2002, Marinova and Jogee 2007). In addition, bars are more common among galaxies with a central starburst (for example Hunt and Malkan 1999), but whether they are more common in galaxies with AGN than those without remains controversial (for example Mulchaey and Regan 1997, Knapen *et al* 2000, Laurikainen *et al* 2004).

4.3. Constraints from stellar populations

Stellar population studies consider mostly visible-light and near-IR colours, not line strengths. As noted in section 3.6.2, combined visible-band and near-IR colours break the age-metallicity degeneracy. However, dust in Sgs dust extinguishes starlight to introduce a new degeneracy. Near-IR colours are relatively immune to dust, but their combination with dust-sensitive visible-band colours ameliorates the age-metallicity degeneracy. To avoid uncertain reddening and contamination by young stars in a superimposed disc, Peletier *et al* (1999) studied nearly edge-on Sgs near $\sim R_e$. They found that the massive bulges of early-type Sgs resemble those of early-type galaxies in the Coma cluster. Specifically, derived ages at $\sim R_e$ scatter by $\lesssim 2$ Gyr around ~ 10 Gyr mean; the absolute age is uncertain from uncertain modelling of near-IR colours. In contrast, smaller bulges in late-type Sgs are bluer, hence younger. As for Egs, where more massive galaxies are older in the mean, the trend of older more massive bulges seems to contradict the prediction that hierarchical mergers form smaller things first. However, one must recall that smaller Sg bulges are probably dominated by secular processes.

When examined spectroscopically, the centres of Sg bulges are found by Proctor and Samsom (2002) to be younger and sparser in α -elements than Egs. Moorthy and Holtzman

(2006) find that large bulges in Sgs resemble massive Egs, both being old and red, with enhanced α -elements. Norris, Sharples and Kuntschner (2006) obtained long-slit spectra to $\sim 2R_e$ in the edge-on S0 galaxy NGC 3115 and find an old, light-weighted mean age for its bulge, while the disc is $\sim 5 - 8$ Gyr old. Given the confused disentanglement of bulge and disc light (Carollo 2005), detailed maps of spectral line indices over a wide range of ‘bulge’/disc contributions are necessary to interpret indices obtained from sparse slits on each galaxy.

Dust restricts our view of the MWg’s inner bulge in visible-band light to ‘Baade’s Windows’, a few sightlines of reduced extinction. In the near-IR, extinction is hugely reduced to reveal a clear triaxial bulge in for example the *COBE* image (Dwek *et al* 1995). Babusiaux and Gilmore (2005) summarize recent work that shows the triaxial bar ending in a possible inner ring whose longest axis is inclined to our l.o.s. by $\sim 22^\circ$. This barred inner bulge, prominent in counts of RC stars, appears to be secular in origin. So, it is surprising and perhaps problematic for the secular evolution picture that stars in Baade’s largest ‘Window’ (at Galactic latitude -4°) are old (Zoccali *et al* 2003, Kuijken and Rich 2002) and have enhanced α -element/Fe abundances (McWilliam and Rich 1994, Barbuy 1999). near-IR imaging spectroscopy of this region will continue to be important for understanding secular bulges in general and the stellar populations of the MWg bulge in particular.

5. Disk Systems

5.1. Luminous Structures

Aside from the aforementioned (pseudo-)bulge population, and a small halo population (as evidenced by metal-poor stars and globular clusters in the MWg and by globular clusters in other galaxies), Sgs are characterized by a prominent disc. Disks are highly dissipated, kinematically cold (rotating) structures, conventionally assumed to have forgotten their progenitors, despite having made only a few tens of rotations since their formation.

The azimuthally averaged SB profiles of discs are exponential in both radial and vertical directions. Many discs are dusty; Driver *et al* (2007) find that 63% of B-band photons released in discs actually escape from these galaxies. As was discussed in section 4, discs suffer non-axisymmetric instabilities that can develop into distortions, predominantly two-armed spirals. The nature of spiral structure is a major subject discussed by, for example Bertin and Lin (1996). Here we concentrate just on issues that relate directly to the formation and evolution of whole galaxies. Specifically, we address

- The existence of thin and thick discs, and the relationship between these apparently distinct components.
- That both ionized and neutral gas can trace Sg discs out to large radii, enabling a more definitive measure of the DM in Sgs than is possible in Egs.
- That Sg discs harbour most of the ongoing SF in the universe, making them important laboratories for studying SF and feedback processes.

5.1.1. Thick disks. Burstein (1979) and Tsikoudi (1979) discovered that the vertical distribution of light in some edge-on galaxies has two exponential components; that with larger scale height is the thick disc. Later, Gilmore and Reid (1983) established a thick disc in the MWg by counting stars by apparent brightness perpendicular to the Galactic plane. Such counts are the convolution of the vertical density function with the very broad stellar luminosity function, so are hard to interpret uniquely because any feature in the luminosity function can be misinterpreted as a feature in the density function. The stellar luminosity

function is broad both because of the mass-luminosity relation for MS stars and because evolved stars range widely in luminosity. The thin disc has $280 - 350$ pc scale height at R_0 and contains most of the baryonic angular momentum. The thick disc has $0.9 - 1.2$ kpc scale height and $\rho_{\text{thick}}/\rho_{\text{thin}} = 0.085^{+0.045}_{-0.065}$ locally (Siegel *et al* 2002). The two discs overlap maximally at $1 - 1.5$ kpc at R_0 ; sparse, faint stars complicate their separation and normalization. Adding colours improves deconvolution prospects.

It is now customary to separate the discs kinematically (for example Carney *et al* 1996), a statistical separation because assigning any star to one disc is uncertain. While, on the whole, stellar metallicity declines from the thin disc to thick to halo (the thin disc has $[\text{Fe}/\text{H}] = -0.7$ to $+0.3$ and the thick has $[\text{Fe}/\text{H}] = -2.2$ to -0.1), there is enough overlap that again only statistical assignment is possible. Reddy *et al* (2006) have observed a sample of ~ 100 thick disc stars, plus fewer stars in the thin disc and halo. It is clear that thick and thin disc stars have strikingly different abundance patterns with little overlap. Specifically, α -elements are relatively enhanced in thick disc stars. In addition, these stars are on average 5 Gyr older.

Suggested origins of the thick disc include a population of dissolved super-star clusters of $\sim 10^6 M_\odot$ (Kroupa 2002), or the infall of satellites (Bekki and Chiba 2000). Reid (2005) argues from the white dwarf luminosity function for a continuous sequence of formation from halo (1% of white dwarfs in the Solar Neighbourhood) to thick (20%) and ultimately thin discs ($\sim 80\%$). The low-luminosity cutoff of the white dwarf cooling curve tells us that the thick disc is $\sim 9.5 \pm 1$ Gyr old, with most of its stars having formed over ~ 1 Gyr; it has $[\text{Mg}/\text{Fe}] > 0.2$ (average in one sample $[\text{Fe}/\text{H}] = -0.49$). Its abundances overlap with those of metal-poor stars in the thin disc. Cooling curves also show that the thin disc began to form many stars 7 – 8 Gyr ago. An age-metallicity relation is evident in the thick disc: $[\text{Fe}/\text{H}]$ increases by ~ 0.5 as the age decreases by ~ 5 Gyr. Reddy *et al* (2006) note that chemical abundance patterns in thick disc stars are distinct from those of stars in dSph's and in the Magellanic Clouds (Shetrone *et al* 2003, Venn *et al* 2004). Thus, the thick disc cannot have been built from mergers of satellites that resemble our surviving current dSph population. Moreover, the distinct abundance patterns of thick and thin discs exclude the possibility that the thick disc could be a dynamically heated thin disc (for example Feltzing 2004).

Bland-Hawthorn and Freeman (2004) propose new instruments on 8-meter aperture telescopes to ‘chemically tag’ stars, to learn the dissipation history of the MWg’s disc. Many elements cannot form in normal stellar evolution, so reflect the chemistry of the progenitor ISM. Stars of expected similar age in the Solar Neighbourhood travel in distinct kinematical ‘moving groups’, suggesting diffusion from common birth sites. Tagging aims to use detailed abundance patterns that are correlated with kinematical streams to track stars with identical abundances back to specific formation sites. In the first test of homogeneity, de Silva *et al* (2006) find negligible intrinsic scatter in heavy element abundances (Zr, Ba, La, Ce, Nd) among Hyades cluster members. The key uncertainty to resolve by study of other clusters is if there is sufficient chemical diversity across the disc to tag all sites uniquely.

Thick discs are ubiquitous. While some early studies found them in early-type Sgs but not in types Sc and later, Yoachim and Dalcanton (2006) have found them in virtually all Sgs sampled. They argue that thick discs result from early mergers of satellites.

5.1.2. Warps and the outer edges of discs. Outer discs are seldom flat. The warp of the MWg was well described by Oort (1961) and extragalactic warps were mapped in HI with early radio interferometers (for example, M33, by Rogstad *et al* 1976). As sensitivity improved, photometry of edge-on (for example García-Ruiz *et al* 2002) Sgs and the rotation curves of highly inclined Sgs (for example Noordermeer *et al* 2007) often show warps in the HI disc beyond the starlight. For example, the HI layer of the MWg thickens monotonically from

R_0 , and beyond $1.5R_0$ warps up to ~ 6 kpc ($\pm 25^\circ$) from the disc plane on one side at 30 kpc radius (the other side is much flatter) (Levine *et al* 2006). Stars follow the warp. Adding a tidal wake in DM to the direct tides from the Magellanic Clouds (Hunter and Toomre 1969), Weinberg and Blitz (2006) explain most of these aspects by tides that resonate in the outer gas disc of the MWg; the $m = 2$ mode is an important constraint that bears on dynamical alternatives to DM and on the shape of the DM halo. The warp is very dynamic, ‘like a flag flapping in the breeze’. The two other Local Group Sgs have prominent, asymmetric warps that may arise from mutual tidal interaction, triaxial DM halos, or accretion of cold gas.

While some discs truncate at ~ 4 exponential scale lengths (deGrijs *et al* 2001), others do not (for example Erwin *et al* 2005). Deep images of NGC 300 (Bland-Hawthorn *et al* 2005) show stars out to 10 radial scale lengths. The discs of a few low-SB galaxies (section 5.4.4) are seen beyond 30 scale lengths, including the prototype of this class, Malin 1, which was found by Barth (2007) to have a normal disc component embedded within the 100 kpc extended component. At such depths, modelling the contamination of starlight by the noise of compact background galaxies is critical; assumptions have been validated against deep galaxy counts from HST.

5.2. Global scaling relations: Mass-Velocity correlation in Sgs

Section 3 showed that constraining Eg internal dynamics is frustratingly difficult. The tensor Virial Theorem shows that massive Egs are flattened by velocity anisotropy not rotation, but we have much to learn about their internals. Considering the homology of global properties of Egs leads to the Fundamental Plane, a global spatio-kinematical correlation.

However, Sgs provide a more favourable situation for deciphering internal dynamics and mass distributions because kinematical tracers in the cold, rotating disc can be measured out to many radial scale lengths. We first consider global scaling relations for Sgs, in analogy with study of Egs that uncovered the FP. In section 5.3, we consider what can be learned about the distributions of light and dark matter from fits to the measured radially variable rotation speed of the disc (the rotation curve, RC). We consider what can be inferred on internal properties from an optimum RC for a single galaxy as opposed to statistical analysis of a sample.

Sg RC’s are measured with emission lines from regions of ionized hydrogen, and with the $\lambda 21$ cm hyperfine spin-flip transition in the ground-state of hydrogen (HI). The latter line is unobscured by dust, and Tully and Fisher (1977) measured its full width integrated across an Sg disc to establish a correlation between the rotation speeds of discs and their B-band luminosities. This Tully-Fisher relation (TFR) is the Sg analogy and precedent of the FP. Its small scatter spurred work on the extragalactic distance scale, galaxy peculiar velocities, and 3D cartography of the local universe. Here, we consider what can be inferred about homology of the global properties of Sgs from the existence of, and small scatter in, the TFR.

The TFR establishes that the disc rotation speed (section 5.3.1) — defined optimally (Verheijen 2001, Noordermeer 2006) by the ‘flat part’ of the RC (although RC’s are not particularly flat, section 5) — correlates with total baryonic mass — as measured originally from the B-band luminosity and now including ISM (McGaugh *et al* 2001) to account for all baryons. The tightest correlation uses K-band luminosity and HI to map baryonic mass while minimising corrections for dust extinction and variable SF to find

$$V_{\text{flat}} \propto L_{K'}^x(\text{baryonic}), \quad (5.1)$$

with (Verheijen 2001, Noordermeer 2006) $x = 4$. One converts to mass with an assumed $\Upsilon_d \equiv [M/L]_d$. Υ_d is constrained by galaxy disc colours (from, for example Bell and de Jong 2001’s grid of stellar population models, to yield $x \sim 3.5$) and by a universal mass-discrepancy

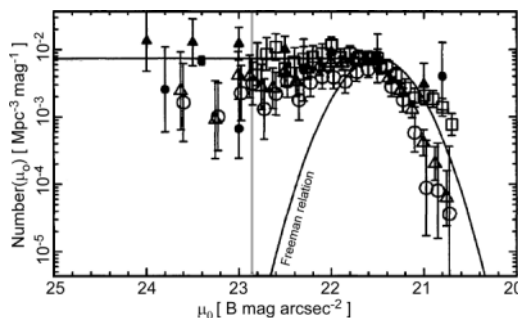


Figure 8. Number density of galaxies versus their central SB μ_0 . LSBg's start slightly to the right of the vertical grey line (the SB of the darkest night sky), and extend to the left where their outer regions can currently be detected down to $\sim 10^{-3}$ of this limit. They can have the same blue luminosity as high SB galaxies, but span a much larger area on the sky. Freeman's selection for high SB galaxies is at right. This figure originally appeared in the Publications of the Astronomical Society of the Pacific (Bothun *et al* 1997; PASP, 109, 745). Copyright 1997, Astronomical Society of the Pacific; reproduced with permission of the Editors.

relation (McGaugh 2004) that is consistent with both colours and standard IMF's. Υ_d seems to vary little among different Sgs (Bell and de Jong 2001, McGaugh 2004).

This tight correlation, coupled with its independence of parameters such as disc size and mean SB (for example Courteau and Rix 1999), shows that the internal dynamics of Sgs do not depend on how baryons are distributed within the potential. Indeed, McGaugh (2005) finds the same behaviour in low-SB galaxies (LSBg's) — those whose average SB is fainter than 40% of the moonless night sky (Bothun *et al* 1997, Impey and Bothun 1997). We adopt the conventional explanation that the fundamental homologies of Sgs arise from ubiquitous DM halos, with baryons simply 'garnishing' the result. An alternative interpretation, MODified Newtonian Dynamics (reviewed by Sanders and McGaugh 2002), proposes to alter Newtonian acceleration at the small value reached toward the visible-band edges of galaxy discs where RC's tend to flatten. Whether or not MOND is real physics, McGaugh (2004) finds that its dependence on acceleration gives the tightest possible description of the discrepancy between baryons and Sg RC's, with scatter arising only from observational uncertainties.

LSBg's are the strongest test of the homology of Sgs through the TFR; they make up to 30% of the gaseous mass of the universe, a negligible fraction of its stellar mass (Driver *et al* 2007), and at least half of the total galaxy population (Figure 8). While mostly detected as gas-rich Sgs in the field (a few being AGN hosts) and in poor, spiral-rich clusters, a few dwarf Egs of low SB have been found. LSBg's challenge Λ -CDM by testing the proposition that baryonic luminosity functions are biased toward high-SB galaxies. To focus on this problem, we now consider Sg RC's and DM in more detail.

5.3. Dark matter content

5.3.1. Evidence from galaxy rotation curves. Contrasting the limited work on DM in Egs, the RC's of gas in the discs of spiral and S0 galaxies provide compelling and direct evidence of DM (for example Broeils 1992, Côté *et al* 2000) because, at sufficient radius, the DM halo mass is responsible for most of the observed rotation. While evidence for important DM originated from RC's of ionized gas in regions of active SF (Rubin *et al* 1978) (HII regions in astronomical jargon), ambiguity lingers (for example Kent 1986) because this gas is generally too faint to see at radii large enough to pin down the DM halo. Specifically, one

cannot distinguish between a maximal-mass stellar disc (i.e., the largest Υ_d whose rotational contribution does not exceed the full RC; Sackett 1997, Palunas and Williams 2000, Weiner *et al* 2001, van Albada and Sancisi 1986) and a sub-maximal disc of lower Υ_d (~ 6.3 according to Bottema 1993's study of stellar velocity dispersions) plus a more massive DM halo (for example Courteau and Rix 1999, Pizagno *et al* 2005). This stellar disc — DM halo mass degeneracy exists from S0 to Sd galaxy types.

While the radial SB profile tracks disc baryons, normalization by Υ_d depends on vagaries of stellar population models and, in discs, on the inevitable presence of dust. Light comes mainly from the upper IMF, in discs from young hot stars that lie near dusty regions. But mass comes from numerous low mass stars. Uncertainty in the universality of an IMF established from the Solar Neighbourhood makes Υ_d hence disc mass uncertain, thence the rest of the RC that is ascribed to the DM halo.

Evidence for DM and its isolation from baryons is compelling in galaxies with the most extended HI, where most maps from radio interferometers such as the VLA and Westerbork Synthesis Radio Telescope show a flat or increasing RC at the largest measurable radius (Rogstad and Shostak 1978, Bosma 1981, Noordermeer *et al* 2007). If the RC spans sufficient radii, one can see variations in the DM/baryon ratio. To establish their individual contributions one can then assume constant Υ_d (questionable) for baryons, and a simple radial distribution of DM, usually an isothermal sphere or the cuspy Navarro *et al* (1997; NFW) halo profile motivated by simulations (section 5.3.1) but with little observational support. In analogy to the relative clarity of bulge-disc decompositions when the observed SB profiles span both bulge- (inner) and disc-dominated (outer) regions, RC's that span radii from where bulge+disc dominate (inner) to where DM halo dominates (outer) allow robust decomposition of baryonic (bulge+disc+gas) and DM components.

RC's from long-slit spectra often jump to 200-300 km s⁻¹ within a few hundred pc radii. The rise to and location of peak rotation correlates with starlight concentration at the nucleus, arguing that bulge mass dominates dynamics (Noordermeer *et al* 2007). RC's often decline 10-20% from the peak before flattening, falling fastest in luminous galaxies. For example, Figure 9a shows Carignan *et al* (2006)'s data and maximal-disc decomposition of M31 within 35 kpc radius, which yields $M_{\text{stellar}} = 2.3 \times 10^{11} M_{\odot}$, $M_{\text{HI}} = 5.0 \times 10^9 M_{\odot}$, and $M_{\text{dm}} = 1.1 \times 10^{11} M_{\odot}$; this galaxy has negligible HI inside 5 kpc radius. DM dominates only at $\gtrsim 30$ kpc.

Indeed, the trend is for high luminosity discs like that of M31 to be maximal in stars, whereas low luminosity galaxies and LSBg's are dominated by DM (for example de Blok and McGaugh 1997, Kranz *et al* 2003). Figure 9b compares mass decompositions of high- and low-SB galaxies. The DM halos (— · —) are very similar despite completely different baryon curves, and DM dominates throughout the LSBg. O'Neil *et al* (2004) map in HI nearly half of 81 LSBg's that have HI mass $> 10^{10} M_{\odot}$; all RC's rise out to the last point measured. LSBg's have the largest Υ_d measured. Their RC's (McGaugh *et al* 2001) are shallower, baryon surface densities are smaller at small radii than those of high SB galaxies, and are best fit with minimal discs.

Flat RC's imply that the DM density must decline no faster than $1/r^2$. Although no erstwhile 'flat' curve has yet been seen to decline at largest radius, the DM profile must eventually steepen to yield finite mass. Halo extents probed with weak gravitational lensing (Hoekstra *et al* 2004) and faint H α emission (Bland-Hawthorn *et al* 1997) both suggest steep DM profiles beyond radii traced by HI. Persic *et al* (1996) argued that all (mostly visible-light from ionized gas) RC's measured by Mathewson *et al* (1992) outside the bulges of Sb-Sd galaxies can be explained by a Universal Rotation Curve (URC) set only by galaxy luminosity. Once a maximal stellar bulge contribution is removed, Noordermeer *et al* (2007)

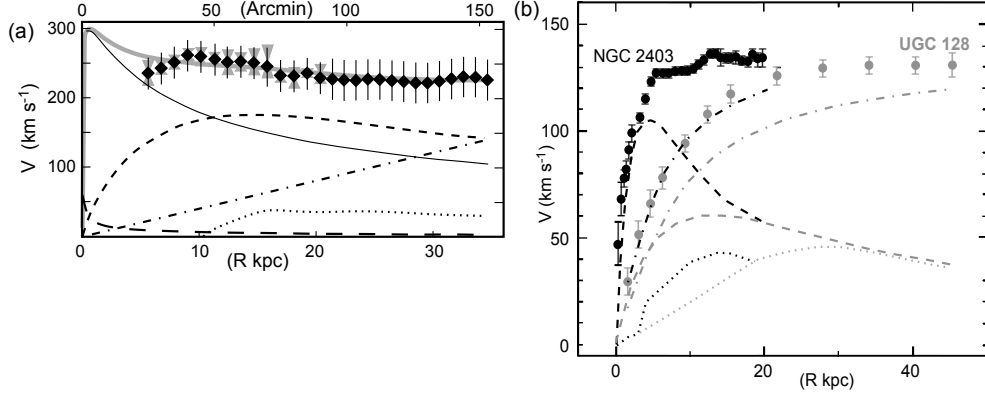


Figure 9. RC decompositions of (a) M31 (grey line and ● Carignan *et al* 2006) with a central $10^8 M_\odot$ supermassive black-hole (— — —) (Bender *et al* 2005) and bulge stars (—); (b) a high SB galaxy (NGC 2403, black lines and ●) and a LSBg (UGC 128, grey lines and ●, adapted from de Blok and McGaugh 1996). Shown in both panels are DM halo (— · —), stars in the disc (— — —) and gas (·····).

find that agreement with the URC is often, but not always, very good. At larger radii, all curves flatten but are too complex to be described by a URC that depends only on galaxy luminosity.

A recent estimate of mass distributions in Sg discs and DM halos from the best RC's is Noordermeer (2006), who analysed HII and HI (out to 20 scale lengths) curves and multicolour surface photometry of 19 S0-Sab galaxies. Rotational velocity totals

$$V_{\text{circ}} = \sqrt{\Upsilon_b V(r) + \Upsilon_d V_d^2(r) + \eta V_{\text{HI}}^2(r) + V_{\text{DM}}^2(r) + V_{\text{pm}}^2(r)} \quad (5.2)$$

with $\eta = 1.4$ accounting for the mass of hydrogen and other abundant elements (gas accounts for only 4% of the total mass), $V_{\text{pm}}(r)$ the contribution of a central point mass (perhaps a supermassive black-hole), and the bulge $(M/L)_b \equiv \Upsilon_b$. Both Υ_d and Υ_b are held constant, despite bluer discs at larger radii that may result from different age/metallicity. Noordermeer *et al* derived mass models for 17 galaxies; Figure 10 shows decompositions for IC 356. A maximal stellar bulge contributes little farther out; the main fitting parameters are the disc scale length and an uncertain Υ_d that yields an uncertain DM halo profile. Nevertheless, the outermost 20 kpc of this RC (and others in his dataset) show clear need for a DM halo. The bottom panel is the best fit without a halo, boosting disc mass density 10 – 50 fold (using an additional baryonic disc component, Pfenninger and Revaz 2005). While residuals are within uncertainties and the fit at $\lesssim 30$ kpc reproduces RC bumps and wiggles of unknown dynamical significance, it is clearly untenable beyond. (He notes that ionization by the metagalactic radiation field may decouple the DM from gas, Bland-Hawthorn *et al* 1997.)

A satisfactory fit uses a DM halo with isothermal profile and a maximal stellar disc (top panel) of larger Υ_d than expected from visible-band colours. An alternative, excellent fit (second panel from top) uses a *minimal* stellar disc with reasonable Υ_d and an NFW DM profile. Hence, definitive assessment of the DM halo profile, and its contribution relative to the disc at maximum rotation speed, is possible only if Υ_d can be found independently. As Noordermeer notes, the most direct constraint on disc mass would come from the velocity dispersion of stars in the outer disc. One would need the very largest telescopes to measure the $\sim 10 \text{ km s}^{-1}$ velocity dispersion of stars at that very low SB. His minimal discs are consistent with simulated DM halos. But, his data show a larger range of central concentrations and his

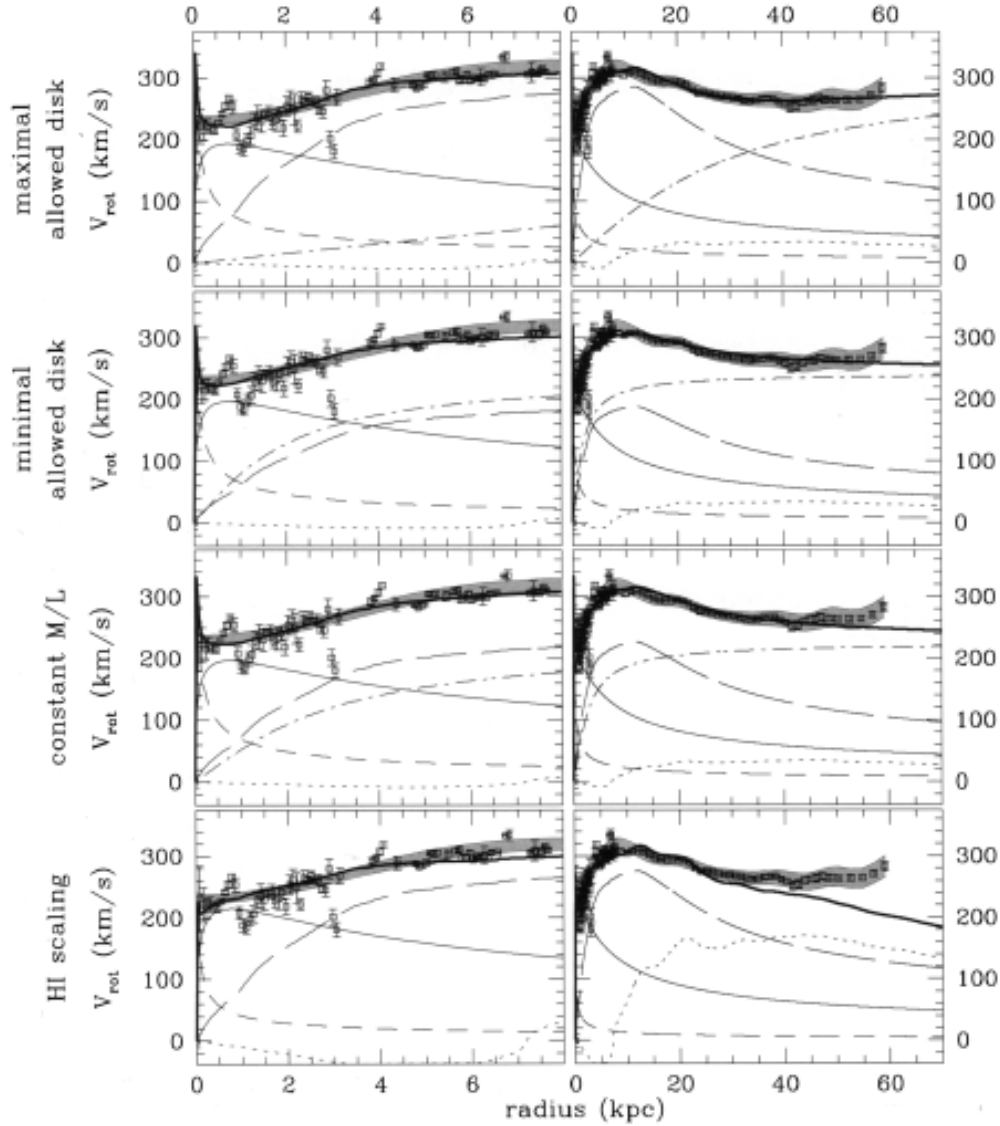


Figure 10. Noordermeer (2006; used with permission)’s mass decompositions of SA galaxy IC 356, using an isothermal DM halo. (Similar results are obtained using an NFW DM profile.) The top two rows assume maximal and minimal mass stellar discs, respectively; the DM contribution is small for the former and large for the latter. the third assumes constant Υ , and the bottom assumes a scaled HI disc of (baryonic) DM and **no** CDM halo. The left column shows the central region with \square showing kinematical long-slit measurements from the ionized gas, the right shows the full RC from HI spectra; the curve width in grey shows the range spanned by the azimuthal average. Bumps and wiggles arise from non-circular motions associated with spiral arms and their resonances. Decompositions are plotted with the following lines: stellar bulge (—) and disc (---), gas disc (·····), central supermassive black-hole (- · -), and DM halo (- - -). Bold (—) is the combined model RC (5.2).

fitted DM halo is therefore more concentrated than simulations find. He finds that core radius increases with scale length, independent of the choice of minimal/maximal discs.

Noordermeer's study shows the advantages of pursuing RC's to the largest possible radius. A complementary approach is to constrain DM and baryonic contributions through statistical analysis of many Sg RC's that range over galaxy mass, SB, and visible-band diameter. Differing contributions of disc or DM halo to the maximum rotation velocity give different patterns of correlated residuals around the mean TFR. For example, Courteau and Rix (1999) studied a sample of luminous, non-barred galaxies. They found little correlation between residuals from their version of the TFR, specifically $\Delta \log V_{2.2}$ versus $\Delta \log R_{\text{exp}}$, with R_{exp} the scale length of the exponential disc and $V_{2.2}$ the velocity at $2.15R_{\text{exp}}$ where rotation peaks in a pure exponential disc. If discs are maximal with no DM halo inside $2.15R_{\text{exp}}$, residuals should correlate with $\Delta \log V_{2.2} / \Delta \log R_e = -0.5$. From this discrepancy and the behaviour of other residuals, they conclude that the DM halo contributes 60% of the total mass within $2.15R_{\text{exp}}$, so even high SB galaxies are not maximal discs at that radius. This largely model independent result has been further reinforced in barred Sgs (Courteau *et al* 2003).

5.3.2. Evidence in the MWg. We are supposedly surrounded by DM, but tracers through the halo are hard to identify. Olling and Merrifield (2001) note another frustration: the local RC of the MWg $\Omega(R_0)$ is uncertain. Errors on 'consensus' values are overwhelmingly systematic, so their propagation is complicated. Thankfully, parallaxes of stars orbiting close to the supermassive black-hole at the centre of the MWg will soon measure R_0 to dispense with some of these irritations.

Microlensing surveys Light variability curves of background stars being microlensed by MWg halo objects have been obtained by the MACHO (Alcock *et al* 2000), OGLE (Udalski *et al* 1994), and EROS (Aubourg *et al* 1993) projects as they stare at the LMC (MACHO) or at the MWg bulge (others). Microlensing measures directly the mass distribution along the l.o.s. As the Earth's orbital motion causes a parallax shift in the observed light curve, the duration of the microlensing event (the time to cross the Einstein ring of angular width $\propto \sqrt{m}$) is set by three parameters that are degenerate observationally: the transverse velocity of the lens to us, the ratio of lens to source distances, and the lens mass m .

Surveys of microlensing toward the LMC have produced somewhat inconsistent results and contradictory interpretations. MACHO toward the LMC finds fewer events than expected if the total MWg halo mass arose solely from compact objects of mass m . Lensing optical depths $\tau \sim 3 \times 10^{-7}$ based on $\sim 13 - 17$ events suggest that 20% of the MWg halo mass is in $0.5^{+0.2}_{-0.3} M_{\odot}$ objects (Alcock *et al* 2000, Bennett *et al* 2005). However, Novati *et al* (2006) proposes that many MACHO events are actually self-lensing by the LMC halo, thereby lowering the estimated DM halo fraction. Moreover, the EROS-2 collaboration finds a microlensing τ only 10% that of MACHO (Jetzer and Novati 2004). Proposed new microlensing surveys will eventually resolve these controversies.

The OGLE/EROS dataset of more than 100 events provides the average τ along several sightlines toward the MWg bulge and the distribution of event durations. Several groups (for example Bissantz and Gerhard 2002) have used these data with far-IR $\lambda 240 \mu\text{m}$ maps of the MWg bulge along sightlines through the stellar bar to argue for a near maximal stellar disc, as found in other Sgs (see section 5.3.1). For example, Bissantz and Gerhard (2002) reproduce the MWg RC out to 5 kpc radius without a DM halo, unsurprising given the small radius probed, and use a near-maximal stellar Y_d . Streaming motions associated with a triaxial bar account naturally for the $> 40\%$ fraction of lensing events that exceed 50 days (Evans and

Belokurov 2002). In short, while microlensing surveys of the bulge have not located a DM halo, they have helped to constrain the inner structure of the MWg.

Stellar densities: disc and halo populations The local mass density can be estimated from the Jeans equation by assuming that the disc potential is steady state, and is separable into r and z motions because vertical variations in mass exceed greatly radial ones. One then obtains

$$\frac{\partial}{\partial z} \left[\frac{1}{v} \frac{\partial (v \overline{v_z^2})}{\partial z} \right] = -4\pi G \rho(R_0, z) \quad (5.3)$$

with $R_0 = 8.0 \pm 0.4$ kpc the solar radius. Measuring the *volume* variations of star density $v(z)$ and $\overline{v_z^2}(z)$ above the MWg plane, yields $\rho(R_0, z)$ (Eisenhauer *et al* 2003). Note that disc star counts must be differentiated three times, once to obtain $v(z)$ and twice in (5.3), strongly amplifying observational errors. Oort (1932) obtained $\rho(R_0, z=0) \approx 0.15 M_\odot \text{ pc}^{-3}$. Hipparcos 65 years later has halved this to $0.076 \pm 0.015 M_\odot \text{ pc}^{-3}$ (Crézé *et al* 1998).

Most halo stellar remnants are old white dwarfs that have cooled below 4000 K. Counts within 1 kpc height yield a halo stellar mass density $\rho_* \sim 0.044 M_\odot \text{ pc}^{-3}$. Deeper surveys are underway, using kinematic cuts to isolate halo and disc white dwarfs, and selecting for bluer colours caused by molecular hydrogen (H_2) absorption in the red; cool white dwarfs have no other absorption lines. From 34 with halo kinematics, Oppenheimer *et al* (2001) conclude that such remnants comprise 2% of the halo mass locally. However, Reid *et al* (2001) argue that this sample mixes at least two kinematical populations, with 20% of ‘halo’ white dwarfs actually in the thick disc (section 3.6.7). Because their disc density declines off the plane, the true white dwarf halo population would be too small to account for much DM. This and the MACHO results exclude a maximal baryon thick disc at least locally.

The disc stellar *column* density within a few scale heights of the plane is better determined because it requires *only* a double differentiation of the star counts

$$\Sigma(< z \text{ kpc}) = -\frac{1}{2\pi G v} \frac{\partial (v \overline{v_z^2})}{\partial z}; \quad (5.4)$$

Oort (1932) found $\sigma(0.7) \sim 90 M_\odot \text{ pc}^{-2}$; Crézé *et al* (1998) review the methodology and efforts to refine his limit. Gould *et al* (1997) use deep HST images to count common disc M-dwarf stars far above the disc plane, normalize these to the locally determined density, and obtain $\sigma(1) = 26 \pm 4 M_\odot \text{ pc}^{-2}$. Olling and Merrifield (2001) obtain $\sigma(1.1) = 35 \pm 10 M_\odot \text{ pc}^{-2}$ and show that the main uncertainty comes from the observationally restricted height into the halo; reliable tracers are sparse, blended with the populations of the discs, and cannot reach the several kpc height required to constrain tightly the DM halo.

Only beyond 30 kpc do we recover DM tracers, now dwarf satellites, GCs, and a smoothly distributed population of evolved RR Lyrae stars beyond 60 kpc (Ivezic *et al* 2004), their kinematics all implying DM (Mateo 1998). Examining these tracers, Olling and Merrifield (2001) use mass models to conclude that the density toward the centre of the MWg’s DM halo is uncertain by a factor of 1000. The relative contributions of stars and DM to the force above the disc plane cannot be separated easily, a ‘vertical disc-halo conspiracy’.

Gas surface density The MWg RC $\Theta(r)$ is constrained by HI data to $\sim 10 \text{ km s}^{-1}$ internal accuracy. However, Lockman (2002) uncovers contradictions in the patchwork of datasets that comprise this curve in textbooks. He cautions us not to take it too seriously, especially between $1 - 2R_0$ where its very uncertain gradient yields a 100% uncertain HI surface density. The curve yields $v_c^2/2\pi G R_0 \sim 226 M_\odot \text{ pc}^{-2}$, ~ 2.5 times that derived from

luminous material within 700 pc of the disc (Olling and Merrifield 1998). Most DM must be beyond this radius, in the halo. The IAU ‘consensus’ rotational speed at the solar radius, $\Theta_0 \equiv \Theta(R_0 = 8.0 \pm 0.4 \text{ kpc}) = 220 \text{ km s}^{-1}$, exceeds slightly Olling and Merrifield (2001)’s best estimate $200 \pm 10 \text{ km s}^{-1}$.

The observed flaring of the HI layer beyond $2R_0$ is a constraint with different dependence on R_0 and Θ_0 . If the ISM is pressurized only by thermally induced turbulent motions in the steady state, hydrostatic equilibrium applies, so

$$\frac{d\sigma_g^2 \rho_g(z)}{dz} = \rho_g(z) K_z(z) \quad (5.5)$$

with gas velocity dispersion $\sigma_g = 9 - 10 \text{ km s}^{-1}$ and mass density $\rho_g(0) \sim 0.042 \text{ M}_\odot \text{ pc}^{-3}$ (Malhotra 1995), and $K_z(z)$ the vertical force per unit mass. Olling and Merrifield estimate that neglecting pressures in the local ISM from the magnetic field (which is frozen to the ionized ISM) and cosmic rays (which are intense only near SF regions) makes less than 10% error in the thickness of the gas layer at R_0 and less at $2R_0$ where there is little SF.

A plausible DM density distribution is an isothermal spheroid

$$\rho_{\text{dm}}(r, z; q) = \rho_0(q) \left(\frac{R_c^2(q)}{R_c^2(q) + r^2 + (z/q)^2} \right) \quad (5.6)$$

with R_c the dark halo’s core radius, q the flattening from the $q = 1$ sphere, and central density ρ_0 absorbs all dependence on q . With R_0 in kpc and Θ_0 in km s^{-1} . Olling and Merrifield (2001) find that

$$\frac{\rho_{\text{dm}}(R_0, \Theta_0)}{10^{-3} \text{ M}_\odot \text{ pc}^{-3}} = \frac{11.5 + 3.8 \times (R_0 - 7.8) \pm 2}{q(26.7 R_0 / \Theta_0)^2} \quad (5.7)$$

The non-linear coupling between these parameters is illustrated in Figure 11.

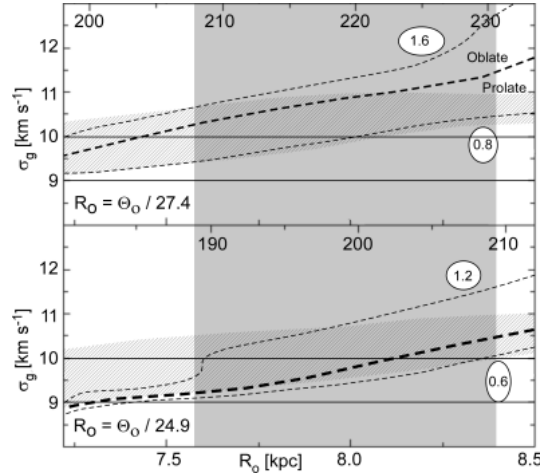


Figure 11. (Simplified from Figure 3 of Olling and Merrifield 2001.) How constraints on MWg scaling constants relate to the gaseous velocity dispersion σ_g and DM halo flattening q . Each panel chooses a Θ_0/R_0 , together spanning their uncertainties; numbers along the top of each panel give the local rotation speed $\Theta_0(R_0)$ in km s^{-1} . Gray shading spans current estimates of R_0 as obtained from parallax measurements of the Galactic Centre (Eisenhauer *et al* 2003). The vertical axis shows a large range over σ_g , with the two solid (—) horizontal lines at Malhotra (1995)’s bounds. Hatched shading spans uncertainties in the column density of the stellar disc. The dark dashed (— —) curve shows the spherical boundary between oblate and prolate shapes of the DM halo and the lighter dashed (- - -) curves show the range.

Implications for the DM halo of the MWg Motions of the Sagittarius dwarf galaxy permit either a prolate or spherical DM halo for the MWg (Ibata *et al* 2001a). However, a ‘Field of Streams’ discovered in the SDSS by Belokurov *et al* (2006) shows that the star stream from this galaxy wraps around the MWg two or three times. Fellhauer *et al* (2006) argue that this coherence requires a close to spherical inner DM halo. (Section 3.6.7 discusses MWg star streams.) Olling and Merrifield (2001) use Malhotra (1995)’s data on the vertical velocity dispersion of HI in the Solar Neighbourhood to constrain flattening of the DM halo; for allowable limits on R_0 and on $\Omega(R_0)$, Figure 11 shows that the MWg halo is spherical or barely oblate and has density in the Solar Neighbourhood $(11 \pm 5) \times 10^{-3} \text{ M}_\odot \text{ pc}^{-3}$ ($0.42 \text{ GeV}/c^2 \text{ cm}^{-3}$). Note that this is within the $\pm 0.015 \text{ M}_\odot \text{ pc}^{-3}$ uncertainty of the stellar density from Hipparcos as mentioned above (Cr  z   *et al* 1998), avoiding substantial DM in the local disc.

5.4. Modes of ongoing SF

One cannot hope to model the evolution of galaxy baryons without understanding SF quantitatively. Unfortunately, dust obscures most SF at visible-light wavelengths. Stars form within an active, multi-phase ISM whose dynamics are only starting to become tractable as long-IR and sub-mm detector arrays have grown in number of pixels, and as 3D hydrodynamical simulations have become more realistic. It now appears that stars tend to form today in the coldest, molecular phase in galaxy discs. SF is stimulated above average levels by any non-axisymmetric instability that sweeps through the disc, usually triggered by close tidal passage or merger of a neighbouring galaxy. This has been quantified observationally (for example Larson and Tinsley 1978), and simulated in detail (Mihos and Hernquist 1996). Examining SF is beyond our scope, see Evans (1999) for example. Here we concentrate on its global aspects that are crucial to modelling galaxy evolution.

5.4.1. Quantification of ongoing SF. How can SF rates be measured accurately in different types of galaxies and environments? SF manifests mainly through the high luminosities of the most massive MS stars formed. So, its unique signature is brief (few Myr), intense UV ionizing radiation from hot stars that is absorbed by gas in the star forming complex. Such HII regions emit hydrogen recombination lines in the visible-band (Balmer series), UV (Lyman), and IR (Paschen and Brackett), forbidden (i.e., electrical non-dipole transition) spectral lines of partially ionized oxygen and other elements, strong UV fluxes, free-free radiation in the radio continuum, and synchrotron radiation from energetic particles produced in SNe (for example Dopita 2005). In fact, because stars form in dusty GMCs, graphitic and siliceous grains quench UV and visible radiation, and recycles much radiation into IR and mm wavebands. The signatures of SF at shorter wavelengths are easily extinguished. Thus IR photometry from space of heated dust is most effective at measuring the true SF rate. Abundances in the gas-phase are set both by astrophysical processes within stars and by chemistry on the surfaces of grains and gas molecules.

The SF rate is set by the energetic photons (and particles) emitted by the most massive stars. Determining the total mass of a star burst therefore requires very uncertain extrapolation of the IMF to the lower masses that constitute most of the event. We encountered this problem in section 5.3.1 when constraining stellar Υ_d . What is observed as ongoing SF is the integrated result of many HII regions of diverse age, mass, size, pressure, and chemical abundance. K-band luminosity measures the integrated light of older stars to establish the efficiency of SF

across a galaxy

$$b = \frac{\text{SF rate}_{\text{present}}}{\langle \text{SF rate}_{\text{past}} \rangle} \quad (5.8)$$

(for example $\sim 5\%$ in isolated dwarf galaxies), with average past SF rate $\langle \text{SF rate}_{\text{past}} \rangle$. $\text{H}\alpha$ flux can be converted into the total flux of ionizing photons using recombination line transition probabilities, determining the rate of forming massive stars; by assuming a form of the IMF, $\text{SF rate}_{\text{present}}$ can be inferred (for example Kennicutt *et al* 1994). More readily observable $[\text{O II}]\lambda\lambda 3726, 3729$ is often a surrogate for $\text{H}\alpha$ in studies of high-redshift galaxies, but is less certain because of the strong temperature dependence of collisionally excited forbidden lines. Kennicutt (1998) reviews determinations of SF rate; he shows that the present SF rate characterized by parameter b varies strongly with Hubble type: b is often < 0.07 for Sa discs, 0.3 for Sb discs, and 1 for Sc discs, all with large dispersions. If the SF rate is taken to decline exponentially, then the e-fold timescale is only ~ 2 Gyr for Sa's, rising to ~ 5 Gyr for Sb's.

Can global SF rates be parametrized for semi-analytic models (and for mixed N-body and gas dynamical models) of galaxy evolution? Are such prescriptions sensitive to environment? For example, tides between galaxies in rich clusters occur more often than in low-density environments. But such encounters at the large velocity dispersion of the cluster will be 'impulsive', with relatively small effect compared to the slow, nearly parabolic encounters of galaxies in groups. Thus, there is a complex trade-off between destabilization from frequent impulses ('harassment', Moore *et al* 1996), and the effects of slow encounters, as modelled for example by Mihos and Hernquist (1996).

Galaxies undergo major bursts of SF, often triggered by tides from a passing companion. A galaxy is star bursting mildly if $b = 2 - 3$ and strongly if $b > 10$. Bursts trigger on a dynamical collapse time: a 10 kpc radius with $10^{10} M_{\odot}$ collapses in 200 Myr; mergers can speed this up. b is measured most effectively by combining SDSS visible-band and *GALEX* space-UV photometry; thus Salim *et al* (2005) find that $\sim 20\%$ of galaxies with $\sim 10^8 M_{\odot}$ to $\sim 5\%$ of those with $\sim 10^{11} M_{\odot}$ have had a significant star burst within the last Gyr. Similarly, Heckman *et al* (2005) have studied 74 nearby galaxies with high far-UV luminosities ($3 - 30 M_{\odot} \text{ yr}^{-1}$ converting into stars) and find that galaxy mass and UV SB correlate inversely. Using the SDSS, Brinchmann *et al* (2004) find that in the local universe most stars form in Sgs that exceed $10^{10} M_{\odot}$ ($\sim 15\%$ being AGN) in regions 100 – 500 pc across, with $\sim 20\%$ in mild star bursts and 3% in intense ones ($b > 10$).

'Passive' SF in high SB galaxies occurs within GMC cores at densities $10^2 - 10^3 \text{ cm}^{-3}$; section 5.4.4 discusses SF in LSBg's. Young stars are dust enshrouded for their first 1-2 Myr, and radiate most ionizing UV within the first ~ 6 Myr, thereafter photo-dissociating their molecular envelopes over ~ 10 Myr. For the next 50 – 100 Myr, burst luminosity increases by only 10% (Leitherer *et al* 1999) as GMC destruction ends SF.

Broad-band colours of visible light age SF crudely (Whitmore and Schweizer 1995). Sensitivity to SF within the last 100 Myr is much improved by adding the *GALEX* NUV and FUV filter bands (Figure 6a). More accurate timing of an ongoing or recently ended burst follows from details of the gaseous emission-line flux ratios. González Delgado *et al* (1999) show that the He I/H β ratio varies with the age of the burst: 0.10 – 0.12 until 4 Myr, then declines steeply below 0.05. The $[\text{O III}]\lambda 5007/\text{H}\beta$ ratio starts near 4.5, declines to ~ 1.2 by 2.5 Myr, and reaches 0.5 by 3 Myr. Near-IR images of hydrogen and helium emission can map SF in dust shrouded discs. In the MWg, emission from transitions between barely bound atomic levels of hydrogen and helium are mapped with radio interferometer arrays across very extinguished parts of GMCs (for example Gordon and Sorochenko 2003).

5.4.2. *Sites of SF and dependence on gas density.* Where do stars form? How does gas surface density set the disc SF rate (Schmidt 1959)? How the rate of ‘passive’ SF depends on the surface density of hydrogen σ_{gas} , and the threshold above which stars can form, was estimated empirically by Kennicutt (1998) from the discs of small-bulge Sgs as

$$\sigma_{\text{SF}} \propto \sigma_{\text{gas}}^{1.4} \quad (5.9)$$

His value agrees with the theoretical ‘disc instability’ criterion for cloud collisions (Toomre 1964), although both Schaye (2004) and Noordermeer (2006) have subsequently mapped ongoing SF in large bulge high SB galaxies below half this value. Silk (1997) gives another estimate

$$\sigma_{\text{SF}} \propto \sigma_{\text{gas}} \Omega_{\text{gas}} \quad (5.10)$$

with Ω_{gas} the orbital speed of the gas clouds. Thilker *et al* (2005) have used *GALEX* to find that stars are forming in some galaxies at 2 – 4 times the radius of the visible-band disc with σ_{gas} below Kennicutt’s thresholds as established at smaller radii. SF rate_{present} depends on the mass fraction in dense gas, whose value is set by turbulence (Elmegreen and Falgarone 1996). Kennicutt (1998) finds

$$\text{SF rate}_{\text{present}} (\text{M}_{\odot} \text{yr}^{-1}) = \frac{L(\text{H}\alpha)}{1.26 \times 10^{34} \text{W}} \quad (5.11)$$

although this misses the, perhaps majority, SF that is extinguished by dust.

Until SF ends with the conversion of all ISM into a warm phase, it can propagate over a $\lesssim 1$ kpc scale (Zhang *et al* 2001); this behaviour is evident in mergers and also as a clear sequence of ‘super-shells’ in the LMC where up to half of the stars form adjacent to such high-pressure HII regions (true also in the MWg).

5.4.3. *Raw material — effects of the state of the ISM.* A major uncertainty in the physical state of the ISM and also the spectrum from hot stars that bathes HII regions is the depletion of elements from the gas phase onto dust grains and polycyclic aromatic hydrocarbon (PAH, a planar-hexagonal carbon lattice built from benzene rings) molecules formed in dusty bubbles around AGB and RGB carbon-rich stars; the size distribution of grains (power-law up to large sizes, declining exponential thereafter); grain evaporation by thermal and perhaps chemical sputtering, grain shattering, and shock heating from SN; the photo-dissociation of H_2 into HII; and PAH photo-dissociation in HII regions. Chemical abundances derived from emission lines show how metals still in the gas phase vary across unobscured HII regions. Most otherwise abundant elements, originally as gas, have depleted onto dust grains; sulphur is one exception, but its emission intensity depends on its temperature, abundance, and level of ionization.

The ISM of the MWg has been reviewed lucidly by Ferrière (2001). Despite ~ 1000 generations of SF in the MWg disc each of ~ 10 Myr duration (Bland-Hawthorn and Freeman 2004), $\sim 10^7 \text{ M}_{\odot}$ of dust and $4 \times 10^9 \text{ M}_{\odot}$ of gas remain to extend SF to 30 kpc radius. Gas subdivides into a cool layer of $\lesssim 0.3$ kpc scale height, and a warm neutral (Lockman) and ionized (Reynolds) layer to 0.6 kpc. Half of its mass is molecular hydrogen (H_2) within cores of GMCs of mass $(0.1 - 2) \times 10^6 \text{ M}_{\odot}$. Velocity crowding of multiple clouds blurs our view. Hence, over the last decade, less edge-on Local Group galaxies mapped in surveys with the BIMA array (N. hemisphere) and 4-meter aperture NANTEN dish (S. hemisphere) have provided our best views. Blitz *et al* (2006) review what these surveys have revealed about the molecular ISM, using emission from the CO molecule as a tracer for the molecular gas. This gas is correlated with HI in that every GMC is found on an HI filament but there are many filaments without CO. CO concentrates toward galaxy centres, with exponential scale lengths comparable to that of the stellar disc. To form stars, HI must clump to become a filamentary

GMC, increasing its surface density up to twenty-fold; this conversion is almost complete in high-pressure, HI deficient regions such as galactic nuclei.

HII regions (Oey and Clarke 1997) and wind bubbles from massive stars break out of the GMC along its density gradient, expanding as the star burst progresses until the gas pressure balances with the lower-density, diffuse ISM. In balance, bubbles then coast in a momentum-conserving state (in which they pass most of their time) that is set by the density of the ISM phase with the largest volume filling factor, until the swept-up shell of ISM fragments by turbulence. In this way and with SN, the GMC core dissolves ~ 1 Myr after SF is triggered (Dopita *et al* 2006). A sequence of cluster ages and correlated CO emission in the LMC shows that a cloud dissolves in $\sim 20 - 25$ Myr (Blitz *et al* 2006; and references therein); HII regions then pass $\sim \frac{1}{4}$ of their lives in a quiescent phase.

The other half of the ISM mass in the MWg is HI in a complex (Cox 2005), multiphase topology that is too uncertain to model evolving SF (although, see the daring simulations of de Avillez 2000). Instead, SF is usually modelled in a uniform pressure and density ISM (for example Dopita 2005; and references therein). In the disc, HI is often optically thick and the uncertain transverse component of its non-circular motions makes uncertain σ_{gas} across much of the MWg (section 5.3.2). Cosmic rays, heavy nuclei accelerated at the shock fronts of expanding supernova bubbles (Ferrière 2001), deposit a third of the energy into the ISM.

Once more than ~ 30 successive SNe have detonated, their individual ‘chimneys’ (in simulations like de Avillez 2000 rising 500 pc above the disc) merge into a ‘superbubble’ (see the review of Veilleux *et al* 2005). These x-ray emitters (thermal spectrum, typically $\sim 10^{5-6} M_{\odot}$) are surrounded by H α -emitting shells up to ~ 1.5 kpc diameter; the LMC has particularly clear examples. If either structure reaches high enough, its thermalized gas redevelops momentum down the pressure gradient, and the flow can entrain and ‘break out’ of the disc as a ‘galactic fountain’: once above ~ 250 pc altitude, it becomes Rayleigh-Taylor unstable from declining ambient densities and eventually drips as descending sheets that pump entrained ISM (including metals, for example Kenney and Koopmann 1999) across a galaxy to smear out primordial abundance gradients. Indeed, at < 1 kpc Lockman (2003) finds dense HI clouds that are too low mass to self-bind gravitationally; they have disc kinematics so were presumably uplifted by correlated SNe and the resulting ‘fountain’. Pidopryhora *et al* (2007) map a superb example ~ 7 kpc distant that reaches ~ 1 kpc into the MWg halo, contains $\sim 10^6 M_{\odot}$ of hydrogen, and has total energy $\sim 10^{53}$ erg.

ISM porosity to supernova blast waves sets the lateral extent of a star burst, the chemical yield, and how efficiently gas is ejected from a potential (see section 6.2). The other important but poorly observed parameter (for example Veilleux *et al* 2005) is the thermalization efficiency ϵ — the fraction of mechanical energy from all SNe not yet radiated before their blast waves merge; it is degraded into turbulence thence entropic heating. ϵ depends on the pressures and scale heights of the various ISM phases, and on the clustering of SNe. Estimated values in star bursts vary widely because of uncertain ISM properties, from $\epsilon = 0.005$ (appropriate for a cloudy ISM with many GMCs and a mix of SNe Ia and II) to $\epsilon = 1$ (appropriate for a galaxy bulge with only SNe Ia). ISM has been incorporated into semi-analytic models of SF by scaling parameters from a three-phase description (see the review by Cox 2005). In this picture, a poorly constrained topology of static cool clouds, warm interface, and pervasive hot phase from intersecting supernova remnants forms and evolves toward $\epsilon = 1$. Blast waves evaporate clouds, and the resulting warm ablata balance cooling from denser regions.

5.4.4. SF in low-SB galaxies. Selections inhibit discovery of LSBg’s, so the correlations discussed previously are only now being extended as targeted searches and surveys like the

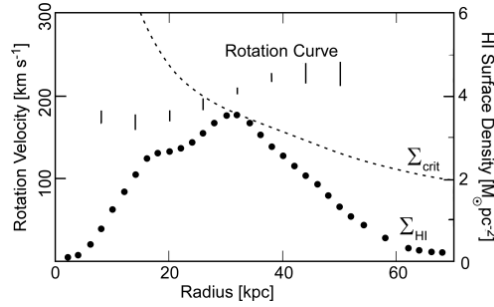


Figure 12. RC (vertical lines) and HI surface density (\bullet) of LSBg UGC 6614. Only at 30 kpc does the surface density rise to meet the critical value of Kennicutt (1998) (\cdots) to form stars. This figure originally appeared in the Publications of the Astronomical Society of the Pacific (Bothun *et al* 1997; PASP, 109, 745). Copyright 1997, Astronomical Society of the Pacific; reproduced with permission of the Editors.

SDSS uncover examples over the full range of galaxy mass, in voids, and on the rims of matter filaments. Most LSBg's have no bulge or bar and little dust. Their blue colours (young stars hence slow, continuous SF), metallicities (0.05 – 0.5 solar, a few > 1), and gas content are less evolved than in high SB galaxies; they are not faded. They are often gas rich yet have HI densities below Kennicutt (1998)'s threshold for SF (for example Figure 12), and sparse CO suggesting no GMCs. Mihos *et al* (1999) model their ISM and conclude that H_2 can be substantial within a few kpc radius but its warmth $\sim 30 - 50$ K inhibits SF. They have extreme HI masses for their blue luminosity, up to 10 times those of high SB galaxies.

5.4.5. Jet and radio-lobe induced SF. Jets can trigger some SF by crushing ISM through direct impact or thermal shocks driven by the expansion of the cocoon/radio lobe (Bicknell *et al* 2000, Fragile *et al* 2004, O'Dea *et al* 2004; and references therein), to explain aligned radio continuum jets and SF (for example Minkowski's Object near NGC 541, Croft *et al* 2006). However, in section 6.2 we discuss suppression of SF ('feedback') by jets.

In summary, disc baryons provide our best opportunity to trace motions in DM dominated regions. We have shown how to measure the SF rate and how it depends on the densities of the diffuse, multi-phase ISM. The diversity of SF complicates interpretation: uncertainties about dust content and IMF make Υ uncertain, complicating isolation of the DM component.

6. Feedback

To parametrize models of the global evolution of star-forming galaxies, the key question is to what extent does energy injected into the surrounding ISM (through UV radiation and winds/SNe from massive stars, production of heavy elements that alter the gas cooling rate, etc) either promote or discourage continuing SF, and redistribute the remaining ISM among its phases? The astrophysical processes responsible are denoted collectively (White and Rees 1978) **feedback**. Because it acts within the underlying galaxy DM potential well, feedback alters the dependences of the SF and chemical evolution rates on galaxy mass.

The crucial rôle of feedback in galaxy evolution appears when modelling the observed luminosity function (LF) of galaxies. In the Λ -CDM scenario, the redshift dependence of the

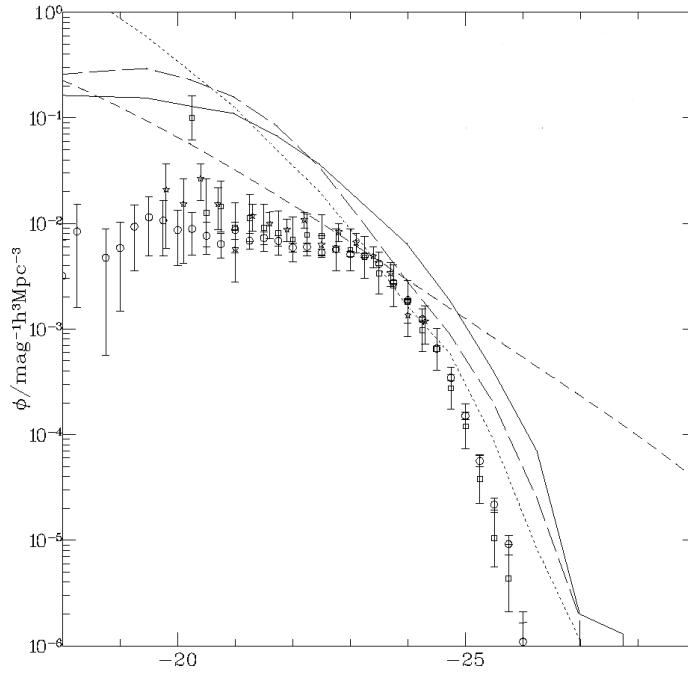


Figure 13. (From Benson *et al* 2003; used with permission). The galaxy luminosity function Φ : the number of galaxies per million pc^3 volume and per magnitude in apparent brightness versus the $-\log$ luminosity of a galaxy (in units of its absolute magnitude, the MWg being -20.7 ± 0.2). Data are \circ , \square , and stars with error bars. Shown are models of the DM halo-mass function without feedback, (diagonal ---), and with various feedbacks mentioned in the text as curves (cooling (.....); photoionization (— — —); merging (—)).

hierarchical assembly of DM halos can be modelled straightforwardly because only gravity operates. The DM halo mass distribution is predicted either from N-body simulations (for example Springel *et al* 2005) or through the analytical framework of the extended Press-Schechter theory (Press and Schechter 1974, Somerville and Primack 1999, Cole, Baugh and Frenk 2000). Because DM halos build by amplifying and merging an initial Gaussian distribution of density fluctuations with a roughly scale-invariant power spectrum that has grown into the non-linear regime, the resulting mass distribution has an exponential high-end and a roughly power-law low-end, as observed (Schechter 1975). However, if one assumes that CDM is modified by a single Υ to yield the baryonic mass distribution (but recall in section 3.4.2 the tight $\Upsilon - \sigma_e$ correlation in Egs), one finds that model and observed LF's disagree. This model LF (--- in Figure 13) has two problems: its low-luminosity slope over-predicts faint galaxies, and it declines steeply at much higher luminosity (mass) than observed. Adding a simple prescription for gas cooling (.....) further over-predicts the faint end because lower mass halos are cooler virially. However, cooling helps at the bright end because gas cools slowly in the massive, virially hot DM halos. Low masses are further suppressed by re-ionization after the first stars and AGN form (— — —). Adding mergers of small galaxies into larger ones (—) further boosts the luminous end. Conversely, this end can be reduced by suppressing cooling in massive Egs with AGN via reheating shocks from their jet-emitting supermassive black holes (Croton *et al* 2006, Bower *et al* 2006), see Appendix B.

While matching the observed galaxy LF, a successful model of galaxy evolution must

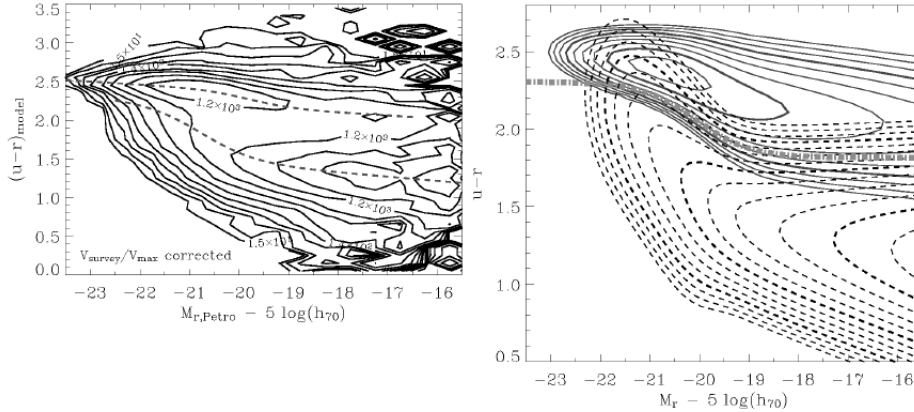


Figure 14. Galaxy colour distribution from SDSS (Baldry *et al* 2004; used with permission). Left: data corrected for incompleteness, with contours showing the number of galaxies. Right: deconvolution into ‘red and dead’ (top (—)) and star forming (bottom (---)) components. Separation is even cleaner when *GALEX* colours are used too.

also reproduce the observed K-band TFR. The zeropoint and slope of the TFR depends on what fraction of baryons cool to form K-luminous stars, while their scale length in the DM halo modifies the RC. Moreover, the Λ -CDM scenario yields a continuum of SF rates in galaxies because it does not inhibit accretion of cold gas that would equip all galaxies with a reservoir from which to form stars today. Yet, the SDSS has quantified (Figure 14) earlier impressions (Kauffman *et al* 2003) that galaxies divide bimodally by colour and mass into a blue sequence of active, SF galaxies and a sequence of ‘red and dead’ early-types that appeared after $z = 1$ (for example, Wake *et al* 2006) with AGN-like emission spectra (Cooper *et al* 2006).

In summary, feedback and other baryonic gas dynamics alter galaxy properties drastically; they ‘quench’ subsequent SF and build up mass by inhibiting flows of energy and angular momentum. Multiple routes in Figure 15 build the red galaxy sequence from the blue; the effectiveness of some depend strongly on environment so are detailed in section 7. Baugh (2006) and Cole, Baugh and Frenk (2000) review comprehensively the parameters of galaxy evolution models. In the following, we merely outline the most important processes.

6.1. Hot and cold modes of gas accretion

Most accretion occurred at high redshift, in two modes: 1) ‘hot’ gas falls into the DM halo at its virial temperature, heating faster than the cooling time in an almost spherical shock. 2) ‘cold’ merging gas cools faster than it is shock heated so collapses first into and then perhaps along a filament (Keres *et al* 2005). In the present epoch, the cold mode is located by simulations (Keres *et al* 2005) at $< 1.8 \times 10^{11} M_{\odot}$ because of lower infall velocities. In contrast, the simulations by Cox *et al* (2004) highlight how the final merger can shift gas from a cold, SF phase to a hot, non-SF one whose ionization depends on the metagalactic UV radiation field, and the uncertain leakage of ionizing photons from the disc and from halo planetary nebulae.

Unclear is how a cold accretor sheds its gravitational energy to bind to a galaxy: through a strong shock near the disc, or gradually enough for pressure gradients to settle gas gently into disc rotation? Fraternali and Binney (2006) consider the extended HI gaseous halos of disc galaxies (for example Swaters, Sancisi and van der Hulst 1997), and conclude from the

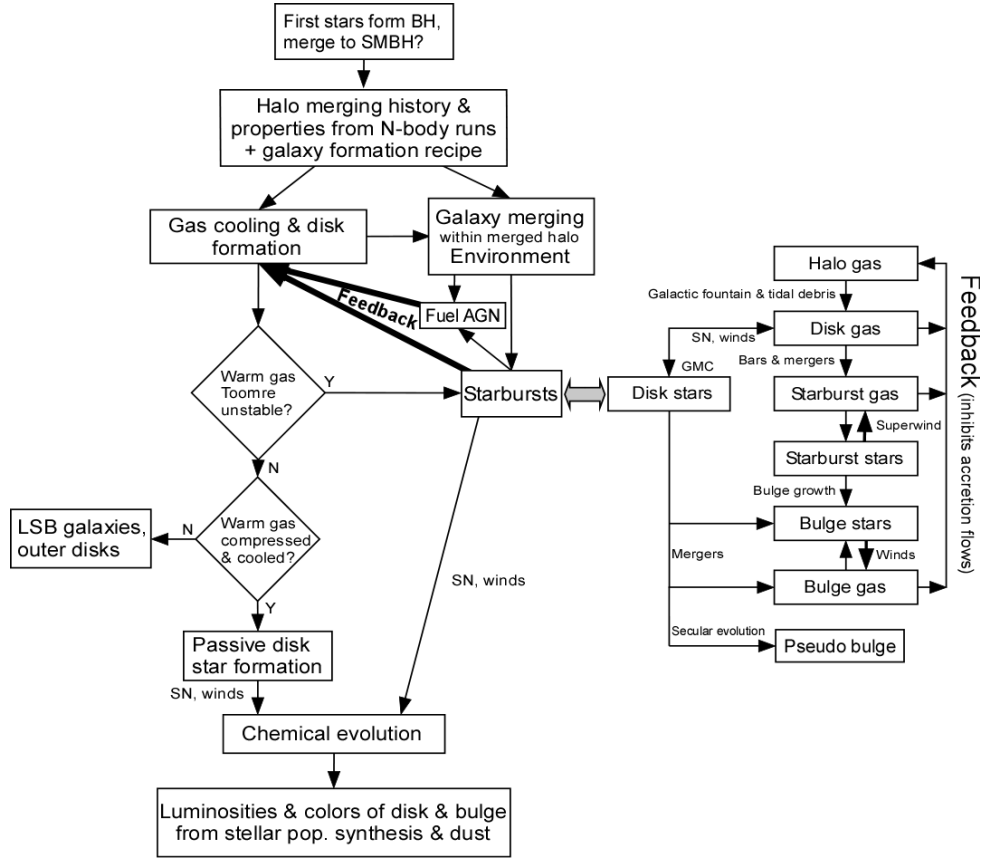


Figure 15. The baryon flows that operate in stellar and AGN feedbacks, and their outcomes. The second box from the top is treated currently with numerical simulations because observational constraints are rudimentary. The other steps are discussed in the text. The flows operating in disc starbursts are highlighted at right.

failure of velocity models to avoid an HI outflow that the galaxy must be accreting from the intergalactic medium. A massive yet isolated HI cloud observed by Minchin *et al* (2005) cannot be tidal debris (Bekki *et al* 2005). Will it become a cold accretion? Metallicity measurements can discriminate provided that ionized gas is present or if there is a background quasar spectrum to absorb its light with neutral gas. Other clues come from the orbits of Local Group metal-poor globular clusters and other satellites of the MWg and M31 (including M33, Koch and Grebel 2006), which are flattened along the major axis of their velocity ellipsoids to indicate non-spherical potentials. No comparable constraints exist for hot accretion at high redshifts, probably because its spherical distribution would be too diffuse to detect with present day x-ray telescopes.

6.2. Mechanisms and scales of feedback

‘Classical’ feedback in a forming, early-type galaxy arises from a large-scale wind established in the ISM by overlapping hot, blast-wave bubbles of supernova explosions (each $\sim 10^{51}$ erg) and stellar winds (each 10% as strong). Recently identified as a major feedback are

the supermassive black-holes in AGN that drive particle jets (section 5.4.5), their shocked cocoons and associated thermal winds over multi-kpc scales and each potentially $\sim 10^{47}$ erg yr $^{-1}$ for up to 0.1 Gyr. Veilleux *et al* (2005) discuss these outflows. An outflow is prominent where directed motion thermalizes, and usually has peak contrast against unrelated emission at a characteristic energy. For example, gas decelerating from v_s km s $^{-1}$ clumps and attains post-shock temperature $T_{ps} = 0.11 v_s / (100 \text{ km s}^{-1}) 10^6$ K that radiates x-rays strongly and is extinguished by photoelectric absorption, not directly by dust. The shocked gas may be further ionized by the AGN's 'ionization cone', although the AGN's radiation field dilutes rapidly with distance. Note that outflows are not the only way to sweep gas from a galaxy; in section 7 we discuss 'ram pressure' stripping as a galaxy passes near the core of a cluster.

The rudimentary treatment of ISM in most semi-analytic models propagates blasts too rapidly through the hot phase, so clouds are crushed quickly to overproduce stars. An ISM more in tune with modern understanding (section 5.4.3) reduces porosity and confines blast bubbles adiabatically to retain most of the supernova energy. For example, Springel and Hernquist (2003) use a two-phase ISM whose components exchange mass by condensing into cool clouds, by evaporating in blast waves, by supernova heating, and are replenished by enriching, energetic wind outflows from massive stars. With this ISM they do predict a quiescent cosmic history of SF that is consistent with red massive, early-type galaxies (Figure 14) and blue, lower mass galaxies (below $10^{12} M_\odot$; Cattaneo *et al* 2006) that cool and form stars.

Gas is also reheated and ejected by AGN. In simulations, an AGN is fuelled by mergers that usually stir star orbits into a bar or other transient structure, which then torque and shock gas into infall. The efficiency of feedback from a supermassive black-hole is set by the nature of the shock propagating medium, and by the efficiency of gravitational energy release $\epsilon_{0.1}$ (units of 10%) such that $E = 2 \times 10^{61} M_{BH,8} \epsilon_{0.1}$ erg (black-hole mass in $10^8 M_\odot$). To remove $M_{g,11}$ (units $10^{11} M_\odot$) from a region of velocity dispersion σ_{200} (units 200 km s $^{-1}$) requires $E = 4 \times 10^{58} M_{g,11} \sigma_{200}^2$ erg. The main accelerator at small radii is radiation pressure $P_{rad}/k \sim 2 \times 10^7 L_{46} r_{kpc}^{-2}$ for ionizing luminosity in units of 10^{46} erg and radius in kpc, respectively; the process is only effective within ~ 10 pc from AGN found locally and ~ 1 kpc in quasars. Coupling AGN ionizing photons to gas is greatly enhanced in dusty clouds (Dopita *et al* 2002).

Springel *et al* (2005) find that a rapid star burst or AGN outflow can end SF abruptly. Cattaneo *et al* (2007) find that intermittent jets are most effective because they do not burrow beyond small radii, hence deposit their energy in the densest gas. Such jets reproduce quantitatively the colour-luminosity, -environment, and -morphology relations of the SDSS dataset, and form quickly the red and dead population of massive early-type galaxies. AGN feedback therefore establishes successfully the bimodality of galaxy types in massive galaxies whose potential wells are too deep for supernova feedback to eject gas, see Figure 14. If accreted gas settled in while a luminous AGN was powered down, the subsequent effects on it after power up could be strong. Dust grains would charge electrically and be expelled, the cold phase removed, the density of the diffuse phase reduced, and the porosity of the ISM to supernova increased from $\sim 10^{-4}$ to 1 over ~ 8 Myr, all by a galaxy-wide ionization front.

6.2.1. Galaxy-scale winds. A galaxy-scale wind regulates SF by heating and entraining ISM into the halo or beyond. Cooling $10^{6-6.5}$ K gas is imaged by the Chandra x-ray Observatory (CXO), the kinematics of gas at $\sim 10^{5.5}$ K are constrained by *FUSE* spectra of OVI $\lambda\lambda 1032 + 1038$ emission at the interface between the cooler bulk cloud and the enveloping ISM (see the simulations of Marcolini *et al* 2005), post-shock gas at 10^4 K is mapped

across outflows in multiple visible-band and near-IR emission lines that can diagnose gas conditions, mid-IR emission from hot dust is mapped by the Spitzer Space Telescope, and HI and molecular lines are mapped with radio interferometers.

M82 and NGC 3079 (Cecil *et al* 2001) provide particularly clear, multi-frequency views of SN-driven, galaxy-winds, while Strickland *et al* (2004) use *CXO* to survey diffuse soft x-ray emission in a sample of edge-on Sgs that spans the full range of SF activity. In thin, star bursting discs, superbubbles blow out vertically, quickly energizing at least the galaxy halo and perhaps the IGM (although, Λ -CDM says that the virial radius of the dark halo of a massive galaxy is ~ 250 kpc and existing x-ray telescopes are too insensitive to track gas to such low SB). In the sharp images of *CXO*, Strickland *et al* isolate diffuse emission from point sources (mostly low-mass x-ray binaries) and find that its SB in edge-on Sgs correlates with the mean SF rate per unit area. Most of the clumped emission in ‘windy’ galaxies turns out to be ambient ISM clouds that have been crushed into pancakes and accelerated by the wind. Depending on how well it resists Kelvin-Helmholtz and Rayleigh-Taylor fluid instabilities, the shredding cloud may approach the wind terminal velocity (Marcolini *et al* 2005).

6.2.2. Ubiquitous AGN and supermassive black-holes.

SWIFT satellite hard x-ray images and HST/ground-based long-slit spectra have uncovered supermassive black-holes in essentially all nearby galaxies with bulges (for example Kormendy and Richstone 1995, Ho *et al* 1997), including some starbursts. The region of gravitational influence is small (at 4 Mpc distance, 1 arcsec for $2.5 \times 10^8 M_\odot$) and in only three cases can one distinguish unambiguously between a dense star cluster and a supermassive black-hole (the MWg, M31 and NGC 4258, see the review of Merritt 2006). A few masses can be obtained from the circular Keplerian motion of compact masers within GMCs that orbit at the outer edge of a nuclear disc (NGC 4258, NGC 4945, and Circinus and NGC 1068 with the complication of severely warped circumnuclear discs), see the review of Lo (2005). More model dependent techniques to bound masses in more distant galaxies are reverberation maps from correlated broad line-AGN continuum flux variations (Peterson 2004), or when the AGN gravitationally lenses a background galaxy into multiple images (Rusin *et al* 2005b).

Our sharpest view of an AGN comes from studies of the supermassive black-hole in the MWg, Sag. A*, which flares in the near-IR/x-ray periodically as it is fuelled from a relativistic orbit (Schödel *et al* 2003). From precise orbits of stars S02 (15.3 ± 0.34 yr period) and S0-16 (comes within 600 Schwarzschild radii, 45 AU), Mouawad *et al* (2005) and Ghez *et al* (2004) derive mass $3.5 \pm 0.3 \times 10^6 M_\odot$, and indeed will soon measure relativity effects such as perigalacticon advance.

Masses of central black holes correlate with host galaxy properties (Marconi and Hunt 2003, Häring and Rix 2004, Ferrarese and Merritt 2000, Gebhardt *et al* 2000)

$$M_{\text{BH},8} = 1.66 \pm 0.24 (\sigma/200)^{4.86 \pm 0.43}, \quad (6.1)$$

(units of $10^8 M_\odot$) implying that the black hole has $\sim 0.3\%$ of a galaxy’s baryon mass. There is some scatter, for example the σ ’s of the MWg (Tremaine *et al* 2002) and M87 (Cappellari *et al* 2006) are too small for their correlation masses, whereas the supermassive black-hole masses in M31 ($1.1 \times 10^8 M_\odot$ Bender *et al* 2005), M32 ($(2.5 \pm 0.5) 10^6 M_\odot$ Verolme *et al* 2002) and Cen A (Krajonović *et al* 2007) are twice those predicted, so perhaps are powered down. Indeed, Revnivtsev *et al* (2004) suggest that the x-ray bright reflector at Sag. B2 ~ 100 pc away indicates that the MWg’s supermassive black-hole emitted $L \sim 1.5 \times 10^{39}$ erg s^{-1} (i.e., 10^6 brighter than today but still only 1% that of a typical AGN) in the band 2-200 keV for at least a decade 300-400 years ago.

Wyithe and Loeb (2003) fit simultaneously the $M_{\text{bh}} - \sigma_{\text{gal}}$ relation and the evolution, shape, and zero-points of the quasar visible-band and x-ray luminosity functions out to $z = 5.5$ in a model wherein a black hole powers up whenever it feasts on gas delivered by a major merger (for example the DM simulations of Volonteri *et al* 2003). They predict $M_{\text{bh}} \propto V_c^5$, independent of redshift. Rafferty *et al* (2006) find cavities in the intracluster medium that were inflated by large radio jets. AGN are able energetically to balance cooling in more than half of the 33 clusters studied. The most powerful example of this heat source was found by McNamara *et al* (2005) in a cluster at $z = 0.22$: a radio source ~ 550 kpc long of $\sim 6 \times 10^{61}$ erg has so heated all the intracluster gas over several Gyr that it has been unable to cool onto the central cD galaxy to form stars.

In summary, both supermassive black-hole and starburst winds are driving outflows that are currently extensive enough to impede mass buildup and to regulate the SF rate over the past Gyr. These and other energy exchange mechanisms in the ISM are sufficiently complex and poorly understood that our treatments of feedback in galaxy evolution remain very uncertain.

7. Rôle of group/cluster environment in galaxy evolution

Section 6 outlined the evolutionary rôles of internally driven feedback. However, evolution can also be stimulated by events within the virialized, hence hot ($\sim 10^8$ K) environment of a galaxy cluster. Ram pressure induced as a galaxy transits this medium can erode the galaxy ISM (Gunn and Gott 1972), and it and stars can be disrupted or merged by tides. We therefore address how environment affects evolution, first considering how cluster studies have constrained DM.

7.1. Dark matter

7.1.1. Evidence from galaxy groups. Evidence for DM around galaxies in groups and bound pairs is slim (Persic *et al* 1996). Mass comparable to that in the stars is estimated from x-ray images (see the review by Mulchaey 2000) by assuming that this gas is in hydrostatic equilibrium in the group's gravitational potential. This is reasonable for the subset of groups with regular, circular shape given the short sound-crossing time compared to local cooling time. The intragroup medium has a significant abundance of heavy elements, usually attributed to galaxy-scale winds (section 6.2.1). But the relative effectiveness of supernova heating in the wind to that of the gravitational potential of the group is uncertain because x-ray images even with *CXO* have not established the efficiency of energy transfer from SNe to the gas because of uncertain gaseous filling factors (Strickland *et al* 2004). The standard assumption is that ongoing galaxy winds are unimportant heat sources and that the hot gas is spherically symmetric. assumptions on gas temperature and metallicity. If the radial x-ray SB follows a King (1962) profile, temperature can also be determined at projected radius. x-ray telescopes before *XMM/Newton* and *CXO* probably missed much hot gas at larger radii, perhaps as much as in the galaxy stars. In fact, isothermal models fit data poorly, arguing again for non-gravitational heating.

7.1.2. Evidence from the cores of galaxy clusters. The same x-ray analysis has been applied to galaxy clusters. Bautz and Arabadjis (2003) derive mass profiles and the central density slopes of 5 clusters with *CXO*; 4/5 are consistent with Λ -CDM predictions. Only the cluster core may be in virial equilibrium. However, in some well studied cases there is clear evidence from x-ray images for transient core heating by shocks from sub-cluster infall (for example Markevitch *et al* 2002). Such structures invalidate the isothermal assumption and may

undermine the assumption of hydrostatic equilibrium. A further complication is a ‘cooling flow’ in the centre of rich clusters and Egs (Fabian 1994) whereby the densest hot gas cools onto the giant cD galaxies at the centres of clusters. In a few cases, central gas is measured to have half the temperature of the cluster mean. On the other hand, accretion rates up to $1000 \text{ M}_{\odot} \text{ yr}^{-1}$ are predicted yet are not seen in soft x-rays. Radio jets from the AGN of the central galaxy appear to be reheating much of the gas and quenching SF (section 6.2.2).

The deflection and focusing of distant background light into multiple images (strong gravitational lensing) maps the distribution of mass across the cores of rich clusters. The pattern of images, for example cluster arcs, has been inverted to constrain the core mass (hence Υ , see section 5 of Bartelmann and Schneider 2001). DM can be mapped in sparser clusters by its shear of background images (weak lensing) to give a lower limit on the cluster core mass. HST’s ACS camera covered three times the area of its predecessor with better spatial sampling, and covered targets out to 1.5 Mpc radii without mosaicing (which complicates derivation of a reliable shear pattern) on more distant clusters (for example Lombardi *et al* 2005). From such a map, Clowe *et al* (2006) find that the DM precedes by 100 kpc a distinct bow shock of hot gas falling through the ‘bullet cluster’ 1E0657-56. The collisionless DM continues unimpeded while the x-ray gas is shocked hence delayed by the collision. DM in more than a dozen clusters has been assessed by lensing and compared to virial and x-ray measures. For example, the various mass estimates of cluster Abell 2218 out to radius 0.4 Mpc are $M_{\text{stars}} = 7.2 \pm 4.3 \times 10^{12} \text{ M}_{\odot}$, $M_{\text{Xray}} = 1.7 \pm 0.5 \times 10^{13} \text{ M}_{\odot}$, and $M_{\text{cluster}} = 3.7 \pm 0.1 \times 10^{14} \text{ M}_{\odot}$, implying that $M_{\text{dm}} = 3.5 \pm 0.2 \times 10^{14} \text{ M}_{\odot}$ and $\Upsilon = 300 \pm 60$ solar (Squires *et al* 1996); there is up to 100 times more mass than visible in stars (Zwicky 1937). The lensing estimate for M_{cluster} agrees with the result from virial analysis of member motions within the potential and from x-ray measurements, although lensing gives the mass inside a cone (along our l.o.s.) and the others give it in a sphere. The discrepancy is largest when the mass filament that contains the cluster is projected along our l.o.s. (Squires *et al* 1996).

Using accurate distances and a least-action infall into the Virgo cluster, Tully and Mohayaee (2004) obtain total mass $1.2 \times 10^{15} \text{ M}_{\odot}$, and $\Upsilon = 850$, seven times higher than that found by the Virial theorem for the core. This analysis is unfeasible for other rich clusters.

7.2. Interactions with the host cluster

7.2.1. Evidence for cluster effect on SF. The morphology–density relation — many more Eg/S0 galaxies than Sgs in rich clusters and *vice versa* in the low-density field (Dressler 1980) — is striking. Does the larger fraction of Egs in clusters arise from their birth (nature) or from evolution induced by the environment (nurture)? A strong clue is the ‘Butcher-Oemler effect’ (Butcher and Oemler 1978): galaxy clusters at moderate redshift ($z \sim 0.3$) have a larger fraction of photometrically blue galaxies than do clusters at low redshift. This finding indicates substantial evolution in cluster galaxies over the past ~ 5 Gyr. Later studies have shown that some blue galaxies have standard Sg emission-line spectra. Others have spectra that are rare at present: very strong Balmer absorption lines without emission. Such resemble the composite of a main-sequence A and a K giant stellar spectrum, so are termed ‘E+A’ or ‘K+A’ galaxies. By inference, their SF ended within the past Gyr, leaving strong Balmer line relics. Their denotation as ‘post-star burst galaxies’ is controversial because it implies that SF ‘ended with a bang’, whereas some researchers argue that SF simply declined suddenly, i.e., was truncated (see section 7.2.2). Balmer lines in some K+A galaxies are so strong that they clearly resulted from a star burst. But only if you catch it immediately after its burst can you distinguish unambiguously a post-burst galaxy from one that is fading with truncated SF.

High-resolution images with HST and ground-based adaptive optics systems clarify that

many blue galaxies are normal Sgs, the spiral fraction in clusters at $z > 0.3$ much exceeds that today (for example Dressler 1997), and there was a larger fraction of interacting galaxies in $z > 0.3$ clusters than at present (for example Lavery and Henry 1988, van Dokkum *et al* 1999). Another clue is that the blue galaxies at higher- z tended to lie outside the cluster core, whereas most galaxies within the core are red even at high z . HST shows that the Eg fraction remains constant at $\sim 15\%$ from $z = 0.5$ clusters to the present; rather, it is the ratio of S0 to spiral galaxies that changes strongly (for example Dressler 1997). Also, Egs at fairly high- z already seem to have a well-defined FP. In summary, the favoured scenario has cluster Egs already concentrated toward the cluster centre at early epochs ($z \sim 2$), whereas the Sg population has evolved rapidly in the past ~ 5 Gyr as they converted to S0 galaxies while falling into the cluster.

More clues emerge from nearby clusters. Caldwell *et al* (1993) and Caldwell and Rose (1997) find early-type galaxies in nearby clusters with K+A ‘post-star burst’ spectra that are generally too weak to classify as true K+A galaxies as defined in higher- z clusters. However, they *do* show the same pattern of enhanced Balmer absorption lines without emission. Also, these K+A galaxies in nearby clusters tend to be rather low luminosity, i.e., they are much smaller galaxies than those seen at higher z . From their major study of Coma cluster galaxies, Poggianti *et al* (2004) argue that evolution in K+A galaxies can be explained by a ‘downsizing’ effect, namely that galaxies terminating SF more recently are less massive. A possible explanation for how downsizing might work is that more massive Sgs form stars more rapidly, so later infalling massive ‘Sgs’ no longer have gas to make the K+A effect.

7.2.2. Conversion of galaxy types. Two routes from spiral to S0 galaxies are proposed: tidal interaction/merger, and interaction between the Sg ISM and the hot (10^8K), diffuse intracuster medium (ICM). Both remove the ISM by stripping or by inducing copious SF. Tidal interaction invokes both processes, i.e., tides can remove significant gas and stars in a roughly equal mass encounter, but the interaction also drains much of the angular momentum from the gas into the galaxy centres to trigger a star burst (Mihos and Hernquist 1996). Tides tend to be strongest in *low* velocity encounters so are disfavoured in galaxy clusters with their very high velocity dispersion. Thus, it has usually been assumed that tides and merging instead make Egs from disc galaxies. Yet there is evidence from high-resolution imaging of many merging/interacting Sgs in $z > 0.3$ clusters (for example van Dokkum *et al* 1999). So, an alternate scenario of ‘galaxy harassment’ (Moore *et al* 1996) destabilizes cluster galaxies by the aggregate of many rapid encounters with other galaxies and/or with the tidal field of the cluster. ‘Wet’ (gas-rich) and ‘dry’ (-poor) mergers are distinguished by the time required to organize the colliders into a spheroid: a dry merger builds a discy bulge quickly by violent relaxation, whereas secular processes (section 4) can randomize the gas discs in a wet merger into a boxy spheroid in $1 - 2$ Gyr.

The alternative to tides is the ‘ram pressure stripping’ hypothesis of Gunn and Gott (1972), who showed that an Sg falling through the hot ICM has its ISM stripped when the ram pressure exceeds the local restoring gravity of the disc. Increasingly detailed numerical models have simulated stripping (for example Abadi *et al* 1999, Schulz and Struck 2001, Vollmer *et al* 2006). Ram stripping should be effective when the density of the ICM is $\gtrsim 10^{-3} \text{ cm}^{-3}$ and the velocity dispersion of galaxies is $\gtrsim 10^3 \text{ km s}^{-1}$, typical of rich clusters. What is seen? First, single-dish studies of many Sgs in nearby clusters show that their global HI content is much depleted compared to field Sgs (Gavazzi 1987, Giovanelli and Haynes 1985, Solanes *et al* 2001). Radio interferometer mapping is more time consuming, so fewer such studies have been made. But they do resolve the HI in cluster Sgs and sometimes show a highly asymmetrical gas distribution, as if the ISM on one side of the galaxy is

being compacted by ram pressure. Sometimes the radio continuum emission trails from the compacted side, and one can see a rim of HII regions concentrated along the squeezed edge (or even ex-planar, as in Kenney and Koopmann 1999), as if SF is being induced along the leading shock where the galaxy ISM plows into the cluster ICM (for example Gavazzi *et al* 1995, Crowl *et al* 2005).

However, in the sparser Pegasus I cluster where the ICM density and galaxy velocity dispersion are low, ram pressure stripping should *not* be occurring. Yet, Levy *et al* (2007) find cases where HI is depleted globally and even concentrates on one side of the galaxy where stars are forming. Other clues exist, but gas stripping from Sgs remains obscure.

Advocating that S0 galaxies in rich clusters like Coma are Sgs with stripped ISM raises questions. Most problematic is that HI-deficient Sgs in clusters appear to have systematically earlier morphological types and higher bulge/disc luminosity ratios than those with normal HI content (for example Dressler 1986). How then can S0 galaxies with their apparently larger bulges be stripped Sgs? Perhaps, as SF exhausts gas, discs fade to increase the bulge/disc flux ratio. Too, Koopmann and Kenney (1998) have shown that the apparent excess of early-type stripped Sgs in clusters is caused at least partially by classification errors in morphology, due to the reduced SF rates. In addition, Caldwell *et al* (1999) have found examples of ‘bulges’ of early-type galaxies in nearby clusters that resolve into star-forming or post-starburst regions (and thus are not ‘bulges’ at all) when viewed in high resolution HST images.

There is also the question of cluster S0 galaxy ages. If these are really stripped Sgs, they would be younger than cluster Egs. Kuntschner (2000) studied both S0 and elliptical galaxies in the nearby Fornax cluster, and concluded that the S0s are indeed younger. On the other hand, Jones *et al* (2000) claim no age difference between these two types in clusters at $z \sim 0.3$. Finally, if S0s form in clusters by stripping Sgs, how do they form in sparse environments?

A related question is: do most of the blue galaxies clustered at $z \sim 0.5$ ‘passivate’ into the non-SF galaxies seen in nearby clusters through a final major burst or by quenching/‘strangling’ their SF as their ISM is stripped? Do data suggest that their SF is simply truncated, or is there a last burst? The answer is controversial. The Canadian Network for Observational Cosmology consortium find no evidence for enhanced SF in $z \sim 0.5$ clusters (Balogh *et al* 1999), so argue that all Sgs in clusters are ‘strangled’. Conversely, the MORPHS consortium (for example Dressler *et al* 2004), find strong Balmer line absorption that cannot be reproduced by the sudden truncation of SF; instead a star burst is required. In nearby clusters, Caldwell *et al* (1999) find that some galaxies at much lower luminosity than the bright star-forming galaxies at higher- z appear to be undergoing, or are seen right after, a star burst. So, although these galaxies are much smaller than big galaxies at higher- z with ‘K+A’ signatures, they do seem to have undergone a last burst of SF not truncation.

In summary, data show that galaxies have evolved in both colour and morphology within clusters since $z \sim 0.5$. Tidal damage and ram pressure stripping drive this evolution, and are being modelled with growing sophistication. However, much remains to be done to develop a coherent picture of rapidly evolving galaxy properties within a cluster.

8. New observational facilities

A basic problem is the strong central concentration of galaxy light, especially in early-types. Light scattered from this point across the field of view of a spectrograph dilutes absorption line strengths. Calibrating this effect is complicated on a modern telescope because its altitude-azimuth mount (adopted universally for rigidity) rotates the field of view during a target track. The diffraction pattern of the mechanical supports of the secondary mirror also rotates

across the field, injecting time-variable structure from the halos of bright sources within and even outside the field. Dusty optics at well ventilated observing sites scatter light over large field angles. Scattered light has an especially insidious effect when photometric profiles are averaged azimuthally, as is often done when one assumes axisymmetry. Diffraction gratings, including modern volume-phase holographic gratings, also scatter light unintentionally as it is dispersed.

Our atmosphere is scattering noise in many ground-based instruments. Shortward of $\lambda 2000$ nm, the incoming wavefront can be ‘de-wrinkled’ with an adaptive optics system to concentrate light. The ongoing retrofit of telescopes with adaptive optics will serve to resolve more distant stellar populations, allowing the powerful CMD discriminators to overcome the degeneracies of spatially integrated spectra. A ‘laser guidestar’ allows an adaptive optics system to operate over almost the entire sky. Pixel SNR scales as the square-root of the exposure time when the signal is Poisson-noise limited. By reducing the size of the sky patch, an adaptive optics system reduces noise twenty-fold or more. Complementing imagery is an IFS to obtain spectra at the limit of the adaptive correction. Hopefully, instruments on the very largest telescope(s) can be optimized to study both faint, distant blobs in the early universe *and* bright, nearby puzzles. It is encouraging that the European Southern Observatory is building the Multi-Unit Spectroscopic Explorer (MUSE)⁺ instrument whose 1 arc-min² field of view will be corrected by adaptive optics. Such instruments could, for example, explore the DM content of an Eg through the kinematics of its dynamically relaxed globular clusters.

Many galaxies are red objects whose dominant stellar population emits mostly in the near-IR. The sky spectrum over $\lambda\lambda 750 - 1500$ nm is covered with numerous and variable OH rotation-vibration band emission. The night sky is dark between, but light scatter within a spectrograph reduces this contrast. Work is therefore underway to modify optical fibres to suppress OH bands prior to wavelength dispersal. No technological barriers have appeared, suggesting that 99% of the bands will be suppressed for another 4 – 8-fold gain in SNR.

After 2013, the James Webb Space Telescope (JWST) will attain high sensitivity longward of $\lambda 1000$ nm with background set by dust emission and scattering (Zodiacal light). JWST will probe dust-shrouded regions of SF in nearby galaxies including the MWg. Missing within 5 years will be UV spectroscopy to assess the full impact of massive stars, so a relatively inexpensive 1.7m aperture World Space Observatory has been discussed. JWST development has already commercialized large-format (4 megapixel), edge-butable. Presently, a single such detector cannot obtain simultaneous J+H-band spectra at the $R = 4000$ required to work in the dark between OH emission lines; a detector mosaic must be used. If one could use OH suppressing fibres, the full near-IR spectrum could be projected at $R = 1500$ onto a single detector and a mosaic could be used to cover more of the sky instead. The ultimate such instrument in the next decade will have a laser adaptive optics system feeding deployable integral-field modules that are coupled by in-line OH sky-suppressing fibres to high-throughput spectrographs.

For brevity, we mention only a few ground facilities under development for use by 2015.

The Large Synoptic Survey Telescope (www.lsst.org) is a 6.5m effective aperture telescope for rapid surveys (15 sec exposures) over 20,000 deg² and $\lambda\lambda 320 - 1060$ nm in six bands. For example, it will map the DM content of galaxy clusters by weak lensing. Very cost-effective liquid mirror telescopes are also being developed for zenith strip surveys.

The 15-telescope Combined Array for Research in Millimeter-Wave Astronomy (CARMA, www.mmarray.org) in the N. hemisphere is currently, and the 64+ telescope Atacama Large Millimeter-Array (ALMA, www.alma.nrao.edu) in the S. hemisphere will

⁺ www.eso.org/instruments/muse

by 2012, provide 0.1 – 1 arcsec resolution (to 0.01 arcsec for ALMA) in millimetre and sub-millimetre CO emission, 5-50 pc resolution at 10 Mpc. GMCs and the dependence of the SF rate on environment can be studied throughout the local universe.

The Global Astrometric Interferometer for Astrophysics (GAIA, www.rssd.esa.int/GAIA) is a 1-meter aperture space telescope optimized for MWg kinematical studies. It will measure parallax distances (up to 90 million stars with better than 5% accuracy and 1 million with 1% accuracy) and proper motions (10^9 stars, 1% of the MWg population) out to a distance of ~ 10 kpc from the Sun for giants and 1 kpc for dwarfs. In addition, it will obtain $R \sim 10^4$ spectroscopy for radial velocities and limited metallicity information near $\lambda 870$ nm for $\sim 10^8$ stars with $V < 17$. GAIA, combined with extensive ground-based followups, will revolutionize our view of the MWg.

This impressive technology will allow us to probe deeper into our nearest extragalactic neighbours, to understand how nature and nurture have combined to present them to us today.

9. Conclusions

The quest to study nearby galaxies across the electromagnetic spectrum has brought these galaxies into sharp panchromatic focus to highlight their energy flows. Contemporary surveys are establishing both the build up of mass/structure and the SF and chemical enrichment histories of the universe. However, while impressive progress has been made, this review shows how precious little we know about chemo-dynamical structures in galaxies and even less about how they operate. Making assumptions about unobserved phase-space dimensions is always perilous.

We have focused on two themes related to nearby galaxies: First, through the behaviour of visible tracers, what can we say about the distribution on their (presumably) dominant DM component? Second, what can we deduce about the internal structure, formation and evolution of their baryonic component?

Despite decades of study, the DM distribution within galaxies remains unclear. As we discussed, SB profiles in Egs are so steep that luminous tracers are too faint for accurate kinematic measurements at the radii where DM is presumed to dominate. Hence, except in the case of Local Group dwarf spheroidal galaxies, the case for DM in individual Egs is still weak, and its distribution is barely constrained. In disc galaxies, DM is better established because RCs are often greatly extended in cold H I. But there remains enough ambiguity in population synthesis models of the visible matter (including uncertainty in the low mass end of the stellar IMF) that the relationship between mass and light is still much debated. Until these issues are resolved, details of the DM will depend on the uncertain visible mass distribution.

In regards to the internal chemo-dynamical structure and evolution of nearby galaxies, we have seen that despite decades of study, how elliptical and spiral galaxies are inter-related is still largely unknown. The most basic questions about the 3-D structure of Egs remain unanswered, beyond the fact that massive ones are primarily flattened by anisotropic velocity dispersion while less massive ones tend to be flattened by rotation. After much controversy over, and gradual refinements in, population synthesis models, we can now extract reliable *light-weighted* mean ages and chemical abundances (of Fe, Mg and perhaps Ca). We are only just beginning to progress to the far more difficult extractions of real SF and chemical enrichment histories of Egs, because these require accurate models of multiple spectral features at many wavelengths.

Our understanding of spiral galaxy bulges is even less certain than Egs, largely because study of the bulge requires identifying and removing the (thin and thick) disc component.

Information about bulge ages still comes largely from photometric colours, not spectroscopic line indices. Recent understanding that pseudo bulges in late type spirals can constrain the hierarchical galaxy formation picture will be capitalized once it is determined how often pseudo bulges arise from purely internal secular evolution and from minor mergers.

The core conclusion of this review is that, if we are to explore galaxy evolution scenarios such as Λ -CDM hierarchical merging by establishing galaxy properties out to high redshift, then we must improve considerably our understanding of nearby, present epoch galaxies. These systems benchmark high-redshift evolutionary studies. The current generation of large ground-based telescopes and space-based facilities contribute by detailing individual galaxies. Surface brightnesses of interest extend deep into the sky-dominated regime, unfortunately yielding only linear improvement in SNR with telescope diameter. Perhaps for this reason, 8-10+ meter telescopes are being used mostly to study high redshift galaxies at large lookback times. However, of equal importance to increased aperture are improved detectors and more reliable control and removal of sky and galaxy light to enable study of galaxies into new regimes of low SB. Too, the development of integral field spectrometers has opened new possibilities for studying galaxy inner regions. In short, there is much to be gained from glimpsing faint wisps within nearby ghosts.

Acknowledgments

GC thanks the Reynolds Foundation for sabbatical support and Director Matthew Colless for hospitality at the Anglo-Australian Observatory during the start and finish of this review. We thank the referee for detailed and constructive comments.

Appendix A. Relevant Astrophysical Processes

Understanding galaxy evolution requires familiarity with two branches of astrophysics: 1) the structure and evolution of stellar tracers of the dynamical and evolutionary state of galaxies, and 2) dynamics of self-gravitating systems. We first review basic stellar evolution theory. We discuss how to simulate the composite light of a coeval population, i.e. a cluster whose stars had the same birth. We introduce methods to compare the integrated spectra of model star clusters of unique age and chemical composition to the observed spectra, aiming to decipher the SF and chemical evolution histories of galaxies. Finally, we give a basic overview of self-gravitating systems, showing how their internal structure can be constrained by their observed light profiles and kinematics. Table A1 lists the acronyms used in this paper.

Appendix A.1. Stellar Evolution

Appendix A.1.1. Basic equations of structure. These have been established for decades, and are discussed in the elegant monograph by Schwarzschild (1958). Throughout the star they characterize local hydrostatic equilibrium, mass conservation, energy production/conservation, and energy transport (either by convection, or radiatively from a temperature gradient). When coupled with an equation of state and after specifying boundary conditions (at a ‘surface’ in wavelength averaged optical depth because the star has no sharp edge), one can integrate numerically the coupled differential equations for internal opacity and energy generation as functions of density, temperature, and chemical composition to obtain a self-consistent model.

Table A1. Acronyms used in this review.

Acronym	Means
AGB	(stellar) asymptotic giant branch
AGN	active galactic nucleus
CBE	(stellar) collisionless Boltzmann equation
CDM	cold, dark matter
CMD	colour-magnitude diagram
CXO	Chandra x-ray Observatory
dIrr	dwarf irregular galaxy
dSph	dwarf spheroidal galaxy
DM	dark matter
Eg	elliptical galaxy
FP	(galaxy) fundamental plane
GMC	giant molecular cloud
HB	(stellar) horizontal branch
HI	neutral hydrogen
HII	ionized hydrogen
HST	Hubble Space Telescope
ICM	inter-cluster medium
IFS	integral-field spectrometer
IMF	(stellar) initial mass function
ISM	inter-stellar medium
IR	infra-red
LF	luminosity function
LSBg	low surface-brightness galaxy
MOND	Modified Newtonian Dynamics
Mpc	million parsecs
MS	(stellar) main sequence
MSTO	(stellar) main sequence turn-off
MWg	Milky Way galaxy
NSAR	non-stellar (chemical) abundance ratios
PA	position angle (on the sky)
RC	(galaxy) rotation curve
RGB	(stellar) red-giant branch
RGC	(stellar) red-giant clump
SB	surface brightness (of unresolved stars)
SDSS	Sloan Digital Sky Survey
SF	star formation
Sg	spiral galaxy
SNe	supernovae
SNR	signal-to-noise ratio
TFR	Tully-Fischer relation
URC	(galaxy) universal rotation curve

Appendix A.1.2. Time dependence. No evolution equation depends explicitly on time. A star evolves because its chemical composition alters gradually through pressure confined fusions. While the ignition condition for each fusion is reached only deep in the star, convection can dredge up fusion products to alter the composition of the inert ‘envelope’. The main driver for evolution is that, as lighter elements fuse into heavier ones in the core, or later in a hot surrounding shell, the mean molecular weight of the gas mix declines to lower pressure. To maintain hydrostatic equilibrium, the core must contract to higher density and temperature. Because fusion rates are extremely sensitive to temperature, core energy release increases thereby expanding and cooling the envelope to brighten and cool the photospheric emission. Chemical composition usually changes very slowly compared to the dynamical timescale. So, the star’s evolution can be traced numerically in time steps, each in hydrostatic equilibrium,

with composition altered by the fusion that occurred during the previous step. On a modern computer it is trivial to evolve a spherical, non-rotating, unmagnetized star. While evolution models have become more sophisticated, Iben (1974; 1991) remain excellent introductions.

Appendix A.1.3. The HR diagram. Before discussing how to derive galaxy ages from precise photometry of their individual stars and/or colours and spectra of their integrated starlight, a brief overview of stellar evolution is useful. Stellar evolution is characterized by the Hertzsprung-Russell diagram, wherein the star's effective temperature T_{eff} , the temperature of the blackbody that radiates the star's photospheric flux per area from radius R_* , is plotted against the bolometric (full spectral) luminosity

$$L_{\text{bol}} = 4\pi R_*^2 \sigma T_{\text{eff}}^4 \quad (\text{A.1})$$

Numerical models evolve a star in this diagram. As hydrogen fusion diminishes and the mean molecular weight of the core increases, the star evolves first to higher L_{bol} and lower T_{eff} in the diagram. In practice it is impossible to measure either of these, because both require absolute spectrophotometry from the UV to the mid-IR. Instead, theory and observation intersect in the so-called colour-magnitude diagram (CMD), wherein the logarithm of a visible-band (or near-IR) flux is plotted vs. a visible-band (or near-IR) colour. Thus one needs accurate transformations between L_{bol} and the particular flux, and between T_{eff}^4 and the particular colour, almost invariably by building a model of the stellar atmosphere to detail its photospheric spectrum.

Chemical composition, aside from providing input to determine a star's age, supplies an independent chronometer for galaxy evolution. Primordial gas from the Big Bang contains only hydrogen (mass fraction X), helium (mass fraction Y), and traces of other light elements (Li, Be, and B). Hence, details of the 'heavy' element content (mass fraction Z) of a star's photosphere gives the degree of prior SF, subsequent evolution, and consequent heavy element production and expulsion (in winds and nova/SNe) prior to the star's birth in the enriched ISM.

Once photospheres of a sequence of stars have been so detailed, correlations with heavy element content may be obtained for grosser spectral features accessible to lower resolution spectroscopy or even filters selected carefully to measure photometric colours. Such data demand less telescope time, enabling study of fainter and more numerous targets in more diverse environments. As datasets are built, the original correlations are extended to less luminous stellar systems, to differing abundances, and to galaxies that have led different lives.

Appendix A.1.4. Colour-magnitude diagrams. Stellar evolution is complex as fusion shifts from core to shell hydrogen burning and then to helium core and eventually shell burning. As the dichotomy between an increasingly hot, dense core and an increasingly extended, cool envelope widens, equations of stellar structure become less certain. We give only the briefest overview of the important phases of stellar evolution. We aim to 1) acquaint the reader with how the various evolutionary phases can date a coeval population of stars, which is crucial to determining SF histories of galaxies, and 2) indicate which aspects of this program are robust and well-understood, and which aspects are uncertain and lead to an uncertain derived age. The underlying question is how accurately can CMDs of Local Group galaxies and the integrated spectrum of starlight from more distant galaxies constrain their SF histories?

The evolutionary track of a star — the path over time in a CMD of a star of given mass and chemical composition — can be computed. By combining tracks over a range of mass but with fixed chemical composition, one can predict the locus of all stars on a CMD at given age, an **isochrone**. An excellent Java Applet, StarPop *, synthesizes the evolution of a stellar

* <http://astro.u-strasbg.fr/~koppen/starpop/StarPop.html>

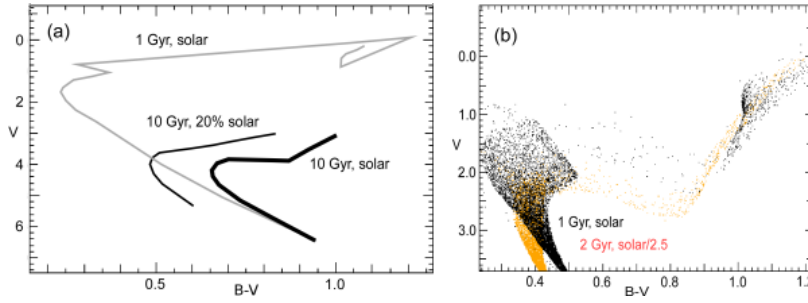


Figure A1. (a) Isochrones from the Geneva Observatory stellar evolution group of 3 different stellar populations, all with identical Salpeter IMF, and plotted for visible light filter bandpasses. The vertical axis is related to log luminosity and the horizontal to colour (hence temperature) with red (cooler) stars at right. The grey isochrone is for a young (1 Gyr old) stellar population of solar chemical composition. Black isochrones are for a 10 Gyr old population of solar composition (thicker) and of heavy element abundance 20% of solar (thinner). (b) StarPop illustrates age-metallicity degeneracy. Two stellar populations arising from a 1 Gyr-long episode of constant SF are overlaid. Black plots a population of solar metallicity that ceased SF 1 Gyr ago. Gray plots a population of roughly half solar metallicity ($Z = 0.008$) that ceased SF 2 Gyr ago. RGB (right) and sub-giant branches largely overlap.

population by sampling isochrones from a specified SF history. Figure A1a illustrates the effects of age and chemical composition on stellar sequences in a CMD.

Appendix A.1.5. Phases of stellar evolution. In Figure A1a, note the long diagonal swath of young, solar chemical composition stars of different mass. Such ‘main sequence’ (MS) stars are fusing hydrogen in cores. In contrast, the older solar-chemical composition population has a short MS because core hydrogen in the most massive stars has exhausted; only lower mass stars are still core fusing hydrogen.

Note the similar evolutionary sequences beyond the MS in both populations. Specifically, atop the MS is the MS turnoff (MSTO), a sequence of stars of increasingly higher mass and luminosity because the higher mass stars are increasingly evolved by depletion of their core hydrogen. In the CMD of a star cluster, the MSTO is our most effective chronometer. Because more massive stars have higher central temperatures and pressures, fusion exhausts fuel more rapidly. Hence, the mass of a star of known chemical composition that is just exhausting core hydrogen can yield a stellar age reliable to ~ 1 Gyr on populations older than 1 – 2 Gyr.

Subsequent, shorter phases of stellar evolution are:

- The red giant branch (RGB), again a sequence of increasing mass, now increasingly evolved. It illustrates the increasing dichotomy between hot core and cool low-density envelope as hydrogen fusion switches to a shell. RGB stars are crude chronometers for populations older than 1 Gyr because stars over a wide mass range funnel into the same part of the CMD. The most important spectral feature for RGB stars is the triplet of singly ionized calcium at ~ 850 nm; it is useful even at lower spectral $R \sim 6000$ that washes out Fe lines. Gallart *et al* (2005) review the uncertainties and conclude that, for a resolved population, the integrated SF rate from its start until ~ 2 Gyr ago can be estimated to within a factor of four from counts of RGB and asymptotic giant branch (AGB, see below) stars, provided that temporarily over-luminous AGB long-period variables are accounted for.
- Subgiant-branch stars, evolving relatively slowly toward the RGB by burning H in a shell

around an inert He core. All stars with MS life longer than 2 Gyr experience this stage.

- The red giant clump (RGC) is the helium core-fusion MS for stars 1-10 Gyr old. Helium ignites suddenly because some electrons in the core became quantum statistically degenerate when the star ascended the RGB, augmenting the normal gas-pressure adjustments for hydrostatic equilibrium with a temperature-insensitive, incompressible ‘lattice’ of electrons. The result is a discontinuous loss of envelope mass and a readjustment of the star’s interior structure. As it ascends the RGB, the star loses mass from its outer envelope at a rate that is ill defined and intractable theoretically, rendering all evolutionary calculations beyond the early part of the RGB uncertain.
- Horizontal branch (HB) stars also burn He in their cores, are low mass, older than 10 Gyr, metal-poor, and have a more extended range of temperatures than RGC stars. RR Lyrae stars are examples, recognized individually because of their variability. The ratio of red to blue HB stars constrains the metallicity and age. Gallart *et al* (2005) review the potential to date ages with the RGC and HB, and conclude that they work reasonably well for 1 – 3 Gyr ages.
- The asymptotic giant branch (AGB). After core helium burning at the RGC, the star burns both helium and hydrogen in separate shells with helium fusion unstable. AGB stars are the most luminous known. They are chronometers for populations older than 1 Gyr, but are less accurate when younger because of uncertain wind-driven mass loss, thermal pulses from helium shell fusion flashes, changes in photospheric abundances from heavy element dredge-up (in particular, carbon), and mixing. Uncertain contribution of AGB stars can alter the fluctuations of the SB of galaxy starlight. If the RGB tip is reached below $\sim 8 M_{\odot}$, the stellar envelope is gone and nuclear burning soon ceases.
- A short, bright ‘planetary’ nebula phase wherein the ejected envelope of a $0.8 - 3 M_{\odot}$ star is photoionized by UV from the exposed helium, carbon or oxygen core. Thereafter, it fades as a low luminosity white dwarf whose core is supported by a barely cooling degenerate electron lattice. White dwarfs are too faint to study beyond the MWg.

The black sequences in Figure A1a show the importance of chemical composition in stellar evolution. Both interior (opacities, hence radiative energy transfer, and mean molecular weight, hence internal pressure) atmospheric (spectral line strengths, hence the emergent radiation field) properties depend on chemical abundances. As a result, the metal-poor (thinner line) MS lies blueward (hotter) of the MS of the metal-rich isochrone (thicker line). The message is that an accurate CMD (from photometric imagery), an independent measure of its chemical composition (from high resolution spectra of the absorption lines in individual stars), and a reliable theoretical isochrone, can be combined to date the population. But, isochrones are only as accurate as stellar evolution, which is uncertain in several areas. There are uncertainties in

- key nuclear reaction rates $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$, $^{14}\text{N}(p, \gamma)^{15}\text{O}$, and even the triple- α reactions are either poorly constrained or have been revised recently to alter evolution timescales (Weiss *et al* 2005). The problem for determining astrophysical cross-sections is the difficulty in producing sufficiently low-energy but high-luminosity accelerator beams.
- opacities that determine radiative energy transport. Widely used opacities are from the Opacity Project (Seaton *et al* 1994, Seaton 1996) and the OPAL (Iglesias and Rodgers 1996) consortium. A larger opacity lowers the H fusion rate and core temperature.
- the chemical composition of even the Sun, the basis of abundance analyses of all other stars. Recent redetermination by Asplund *et al* (2005) indicates that the abundance of O and other major elements must be reduced ~ 1.5 -fold, a controversial result because

it worsens agreement between theory and helioseismology spectra of the solar interior (for example Bahcall *et al* 2005, Antia and Basu 2006). Increasing helium abundance Y increases T_{eff} and L_{bol} and decreases the width of the MS.

- convection, which does not seem to have even a mid-term solution. We know so little about it that models parametrize it by a convective scale height, the typical distance travelled by a convective cell in units of the local pressure scale height. Because convection transports energy in the outer layers of the Sun, its scale height sets the solar radius. So, this scale is fixed *ab initio* in a 4.6 Gyr model to reproduce the solar radius, and its value is assumed to apply universally to stars whatever their internal structure, increasing the core mass. Our ignorance is further compounded by convective overshoot: the scale is increased to account for the inertia of bubbles that turn around not where their acceleration reaches zero but beyond, at zero velocity. Overshoot is clearly highly uncertain, but appears to be necessary to match model isochrones with the CMDs of intermediate-age star clusters (for example Kozhurina-Platais *et al* 1997). Finally, convection plays a huge rôle by intermixing layers that are otherwise stratified into different chemical composition. For example, deep mixing transports fresh hydrogen into the core, prolonging a star's life. Models for intermediate-mass stars are compared by Dominguez *et al* (1999), who find 10% variations in H-burning lifetimes and 30% for He core-burning.
- handling mass loss. As a star ascends the RGB, its envelope distends and becomes less bound. An AGB star loses mass through either a gradual wind or sudden thermal pulses. The empirical Reimer's Law (Schröder and Cuntz 2005) estimates gradual mass loss by a wind from an RGB star, but data are hardly reliable enough to predict the star's mass when it ignites helium in its core. Thus, all evolutionary calculations become more uncertain once the star ascends the RGB. Specifically, the location of a star in the HR diagram during its core helium burning, HB phase depends on the envelope mass (whereas the mass of the helium core, hence the star's luminosity, is relatively insensitive to initial mass): stars with lighter envelopes have hotter T_{eff} , so are bluer. Uncertain mass loss confounds prediction of a star's position along the HB. Hence, isochrones that include HB and post-HB phases have made basic assumptions about mass loss that usually force the morphology of the HB to agree with that observed in star clusters.
- element diffusion from gravitational settling in the outer, radiative layers that can mask the star's true chemical composition. Helium diffusion in the core will alter the star's lifetime. An independent check on its efficiency comes from other elements, Staniero (1997) suggesting those not burning at H-burning temperatures, Fe or Ca, as candidates and both measurable in Local Group galaxies down to the MSTO with large telescopes.
- the unobserved stellar rotation that augments hydrostatic equilibrium and breaks the spherical symmetry of the star to greatly complicate models. Rotation is ignored when many models are generated for an isochrone. Rotation alters deep mixing hence chemical abundances.
- Coulomb effects that soften the equation of state. They reduce MSTO ages by 10-15%, and ages derived from B- and V-band colour differences by 40-50%.

These uncertainties grow at more advanced evolutionary stages as the internal structure of the star becomes more extreme. Hence, the most robust estimate of the age of a star cluster remains the T_{eff} and/or luminosity of the MSTO, as determined from a CMD. As we will see, when individual stars cannot be resolved, one can still infer MSTO characteristics by deciphering the integrated spectrum of galaxy starlight.

Appendix A.1.6. Sensitivity to basic parameters. Broad bandpasses bluer than V cannot find T_{eff} accurately because the spectrum is blanketed by a forest of absorption lines whose strengths vary with chemical abundances and surface gravity $g = GM/R^2$. Larger telescopes and more sensitive detectors have supplanted photometry with spectral determination of stellar parameters. For $5000 \text{ K} \leq T_{\text{eff}} \leq 8000 \text{ K}$, Balmer line strengths yield accuracy $\pm 50 - 80 \text{ K}$. Surface gravity can be established to a multiplicative accuracy of 1.5–2.5 by its pressure modification of some line profiles (mainly Fe, Ca I, Mg Ib triplet at $\lambda 520 \text{ nm}$, and CN bands at $\lambda 380 + \lambda 410 \text{ nm}$) (for example Fuhrmann *et al* 1997). Lebreton (2001) discusses how star diameters derived from interferometric measurements have calibrated to internal accuracy $\sim 1\%$ a combined empirical/model atmosphere Infrared Flux method to deliver T_{eff} for A-F dwarf and giant stars (Blackwell and Lynas-Gray 1998) and, an SB method with comparable accuracy (DiBenedetto 1998). Recent work (for example Bessell *et al* 1998) improved treatments of atomic/molecular line blanketing in O-K and A-M stars.

Appendix A.1.7. Sensitivities to variations in Y and Z. Lebreton (2001) reviews measured chemical abundances; Wallerstein *et al* (1997) review the derived patterns. Abundance estimates are less accurate for hot and young stars because most of their ionization structure is in the UV, and because spectral features are blurred by rapid stellar rotation. Elements to study include

- O, Mg: products of SNe II that probe contributions from $> 20 M_{\odot}$ and trace chemical mixing in globular clusters.
- Si, Ca, Ti, Cr: also products of SNe II but less strongly weighted toward more massive progenitor stars. Combined with O and Mg, these ‘ α –elements’ (even-proton number nuclei) probe the distribution of initial masses (the so-called initial mass function, IMF) (McWilliam 1997).
- Mn, Co: from SNe II, with yields depending on metallicity of the progenitor star.
- Eu: indicate rapid (r-process) neutron pickup (compared to the rate of β -decay) nucleosynthesis in SNe II.
- Y, Zr, Ba, La: indicate slow (s-process) neutron pickup nucleosynthesis in lower-mass AGB stars.
- Fe, Ni: track metallicity from SNe Ia. Fe also provides parameters such as T_{eff} and g .

Detailed abundances constrain the SF history. α -elements come preferentially from hydrostatic burning in massive stars before their SNe II make r -process elements. Figure A2 (Fig. 1 of Bland-Hawthorn and Freeman 2004) shows how successive generations of low-mass stars enrich by SNe II. Roughly 1 Gyr after the star burst, white dwarf deflagration/detonation SNe Ia (from mass accretion in a binary star system) make more Fe-peak elements. An observed overabundance of α -elements compared to Fe-peak elements implies a brief or recent SF episode (Trager *et al* 2000a, Trager 2004); their detailed abundances depend on the mass of the supernova progenitor. During this time, AGB stars make odd-proton nuclei in the s -process. The ratio of α /Fe-peak element abundance decreases as a star burst progresses. So the relative abundances of r - to s -process elements gives the relative importance of SNe Ia to AGB stars, hence the relative yields of higher and lower mass stars.

Appendix A.1.8. Uncertain isochrones. The pace of stellar evolution is exponential with mass. To combine tracks to form an isochrone, one must interpolate in the $(\log M, \log X)$ plane

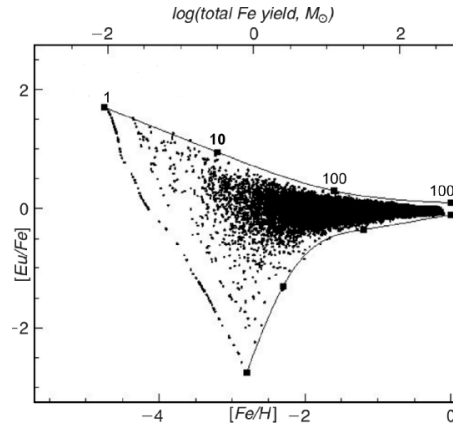


Figure A2. (From Bland-Hawthorn and Freeman 2004; used with permission of Publications of the Astronomical Society of Australia, the authors and the publisher CSIRO.) How successive generations of low-mass stars are enriched by just SNe II, as traced by europium production (r-process). The initial wide scatter in abundances at left converges to a universal value at right. The stars are formed with a Salpeter IMF and the element yields are from Tsujimoto and Shigeyama (1998). The upper bound comes from lower-mass SN, the lower from higher-mass. From left to right, ■ indicates the number of prior enrichments: 1, 10, 100, 1000. Just 10 enrichments from high-mass SNe suffice to enrich a cloud to $[Fe/H] = -2$.

(X is the core hydrogen mass abundance) by assuming that stellar evolution is continuous in this plane (it is not), and that tracks of stars of slightly different mass are always proximate in the CMD (also untrue in some mass ranges). Isochrones may be interpolated physically only when homology relations can scale stellar properties. Homology is broken by mass loss and when core processes switch drastically over a small mass range such as the dominance of CNO over proton-proton fusion. Interpolation to intermediate masses then yields invalid tracks through unphysical parts of the CMD. Homology relations anyway predict only indirectly observable interior — not observable photospheric — properties. Finally, all models ignore important evolutionary influences such as rotation, binary mass transfer, and magnetic fields.

Isochrones differ slightly depending on how RGB and HB convections are handled (Gallart *et al* 2005). For example Geneva (http://obswww.unige.ch/~mowlavi/evol/stev_database.html) and Padova (<http://pleiadi.pd.astro.it/>) group zero-age isochrones span observations of both halo and disc stars only if temperatures are set 130–250 K hotter than observed.

Appendix A.1.9. Mass functions. Stellar $M/L \equiv Y$ is found empirically to vary from 4×10^{-5} (O5 star) to 34 (M5). This mass-luminosity relation is fundamental, yet its measurement is limited to those stars in suitable binary orbits that are also near enough to us to have a precisely measurable parallax; a 10% parallax error produces a 5% error in T_{eff} .

Models must specify a stellar birth-rate. The usual simplification forms stars in one burst, so have a single age, metallicity, and initial mass function (IMF) that together define a so-called simple stellar population for the entire region observed. The IMF is the probability of finding a star of given mass; it thus also sets the mass fraction of the population that is returned to the ISM by evolved stars. As a distribution, its use implies an unavoidable Poissonian sampling variance (see Cerviño *et al* 2002; for a correction procedure); for example, the predicted colours of a $10^5 M_{\odot}$ cluster (which is ~ 6 times the mass of the most massive

OB star association known in the MWg, Cygnus OB2) disperse by 3-10% depending on the assumed IMF and on whether or not a few rare but luminous stars are included.

IMF's are based on censuses of nearby SF regions. However, these regions make predominantly low mass stars, are subject to variance, and are complicated observationally by unresolved binary stars. The Salpeter (1955) power-law IMF $\zeta(m) \propto m^{-\alpha}$ with $\alpha = 1.7$ is standard and has been assumed in most studies. However, Kroupa (2001) summarizes evidence that this IMF predicts more low-mass stars than are observed in MWg and LMC open star clusters; he provides an alternative, multiple-part IMF from fits to those regions. The mean stellar mass with this IMF is $0.36 M_{\odot}$; half of the mass is in stars $0.01 \leq m \leq 1 M_{\odot}$ and the other half is in $1 \leq m \leq 50 M_{\odot}$. The variances arise mostly from unresolved binaries.

Bruzual and Charlot (2003) show that Υ derived from different plausible IMF's varies by multiplicative factor ~ 1.5 for up to 1 Gyr after a star burst, increasing to multiplicative factor ~ 2.5 for ages up to 10 Gyr. Another uncertainty on the low end of the IMF is that the current census with mass $< 0.5 M_{\odot}$ is incomplete beyond 5 pc (Henry *et al* 1997). The high end of the IMF (especially for Wolf-Rayet stars) has uncertain normalization and highest variance from observations (see for example Fig. 8 of Bruzual and Charlot 2003); this is irrelevant for models of early-type galaxies that have not formed many stars recently. In their high-SNR study of 25 such galaxies (see section 3.4), Cappellari *et al* (2006) compare the dynamically determined Υ with the values implied from population synthesis (Vazdekis *et al* 1996) to fit the observed absorption strengths. They obtain better agreement between the two methods across their sample with the Kroupa (2001) IMF; its fewer low-mass stars compared to Salpeter reduces the derived Υ and hastens the subsequent evolution of the star cluster.

Appendix A.1.10. Building a model galaxy. Having considered the evolution of a single star (evolutionary track in the HR diagram) and of a stellar system (isochrone), we can now model a simple stellar population, i.e., a stellar system that is co-eval with unique chemical composition. A grid of SSP's, ranging in age and chemical composition, forms the basis for comparison with real galaxies. To create the spectrum of an SSP, one selects an isochrone of desired age and metallicity and an appropriate IMF. Next, at each point in the HR diagram, the number of stars (weighted according to the isochrone and the IMF) is calculated, and a stellar spectrum is selected with the appropriate T_{eff} , $\log g$, and metallicity. Finally, the spectrum is weighted by the number of stars at that point and by luminosity, and combined with those at all other points in the diagram to synthesize the spectrum of the particular SSP. Bruzual and Charlot (2003) detail this procedure. Naturally, its success depends on having a spectral database that covers a wide range in T_{eff} , $\log g$, and chemical composition. Because of the difficulty of modelling the emergent spectra of stars with millions of atomic and molecular transitions, it has been traditional to use 'libraries' of uniform-quality spectra of many stars whose T_{eff} , $\log g$, and metallicity has been determined from a high dispersion abundance analysis. Recent libraries, each comprising spectra of ~ 1000 stars and ranging widely over atmospheric parameters, are

- The Indo-US library of coude-feed stellar spectra (Valdes *et al* 2004) at www.noao.edu/cflib, covers $\lambda\lambda 346 - 946$ nm at resolution $\sim \lambda 0.1$ nm FWHM for 885 stars.
- ELODIE (Moultaka *et al* 2004) spans $\lambda\lambda 400 - 680$ nm at high dispersion. Based on stars with well-determined atmospheric parameters, version <http://atlas.obs-hp.fr/elodie> has $> 10,000$ spectra.
- MILES (Sanchez-Blazquez *et al* 2006) contains spectra at resolution $\lambda 0.23$ nm FWHM and covers $\lambda\lambda 352 - 750$ nm.

Empirical libraries have limits. It is difficult to find stars that range over sufficient T_{eff} , $\log g$, and metallicity whose atmospheric parameters are known from fundamental high-dispersion spectral analysis. This is particularly true for hot stars, because only metal-rich examples exist in the Solar Neighbourhood and it is now clear that the mean abundance ratios of stars in Egs are non-solar (see section 3.6). Establishing atmospheric parameters for stars in spectral libraries is also problematic, because the high dispersion analyses were made by different investigators using different model atmospheres and T_{eff} scales from different photometric colours.

Models of stellar atmospheres have increased in sophistication through great increases in both computational power and the number of included atomic and molecular transitions. Hence, using a library of synthetic stellar spectra is increasingly attractive. Theoretical stellar spectra can be calculated for any chemical prescription. For full consistency, one should first model an atmosphere based on the assumed chemical composition, and then calculate the detailed emergent spectrum. While calculating the model is very time consuming, use of theoretical spectra is now widespread especially for constructing the hot stars in young stellar populations. An example is the library of 1654 theoretical spectra, sampled at $\lambda 0.03$ nm and covering $\lambda\lambda 300 - 700$ nm, that is described in Martins *et al* (2005)[‡].

Two approaches compare composite spectra of SSP's, gridded by age and metallicity, to the observed spectrum to determine which combination of SSP's matches best.

- (i) Focus on absorption lines in the model grids with particular sensitivity to age and/or metallicity, then find their best match. The preeminent example is the Lick system of spectral indices (for example Faber *et al* 1985). This approach has several attractions:
 - It is transparent to which specific feature(s) determine age and metallicity.
 - It is simple to compute a two-index diagram that separates age and metallicity effects (for example $H\beta$ versus the Fe5270 index), plot the grids that connect the various SSP models, overplot galaxy measurements, and then interpolate ages and metallicities for a large galaxy sample.
 - With model atmospheres, one can probe precisely which transitions of various elements form the index, either through the central passband or through the two continuum bands that straddle it in wavelength.
- (ii) Use χ^2 minimization to fit the entire spectrum to models. This approach is problematic because a typical spectrum contains more than 1000 independent pixels, so computation is intensive and impractical for the numerous ($\sim 10^6$) spectra from an extensive survey such as the SDSS. Moreover, spectral information is highly redundant, with many features duplicating the overall degenerate sensitivity to age and metallicity. If the SNR is low, co-adding many redundant features is advantageous, although populations would then not likely be isolated. If the SNR is high, information in the spectrum can usually be reduced to a few parameters that are chosen judiciously. Various strategies are possible:
 - Pre-select line strength indices, as in 1) above.
 - Use principal component analysis (Madgwick *et al* for example 2003, Ferreras *et al* for example 2006) to find the minimum eigenvector set that contains most of the spectral information. Solving for the eigenvalues that characterize a given observed spectrum yields enormous compression. However, spectral evolution of stellar populations is very non-linear in time. Thus, for a set of emission-free spectra of different ages (to avoid the vexing emission-line 'contamination' of absorption

[‡] www.astro.iag.usp.br/~lucimara/library.htm

features), the galaxy spectrum is synthesized by combining a very young with a very old population that may be irrelevant to that galaxy's SF history.

- Compress spectra through a set of weighting vectors to match a set of realistic population parameters (for example a dozen ages spaced logarithmically, each with a metallicity). Panter *et al* (2003) have used such compression (acronym MOPED) to extract SF histories from huge numbers of SDSS spectra. Of concern is that this approach fits data with a highly compressed χ^2 minimization that may fail to duplicate faithfully the observed spectrum. While the same criticism can be levelled at the spectral index approach, mismatches between models and data are there more obvious. For example, plotting the age-sensitive Lick H β index against the Fe-sensitive Fe5270 index on a model grid with various ages and metallicities but using solar abundance ratios, yields a single metallicity. For massive Egs, plotting H β against Lick Mg b indices (Mg b measures primarily a Mg-sensitive feature) yields a larger metallicity. Such results have indicated clear non-solar abundance ratios (NSAR) in Egs (for example Worthey 2004, Trager *et al* 2000a, Peletier 1989), requiring consistent NSAR models (Thomas *et al* 2003). Such transparent disagreement between model and data led to important insights on the SF/chemical evolution histories of galaxies; it is far less clear with the MOPED and principal component techniques.

Appendix A.2. Gravitational stellar dynamics

Goals of gravitational dynamics are to

- (i) Understand the phase-space distribution of stars in galaxies. The observed 2D projected stellar motions and SB distribution constrain the 3D shape and motions in the potential. The stars, with the inferred DM, establish the gravitational potential of the galaxy. (Gas is generally negligible in gravity, simply acting as a tracer that may not be in equilibrium.)
- (ii) Establish the stability and secular evolution of the stellar and DM distributions in galaxies. How do these two interact over time in 'isolated' galaxies?
- (iii) Establish the consequences of gravitational interactions between galaxies.

One must recognize that a self-gravitating system differs greatly in physics from a charged plasma, perhaps making gravitational dynamics seem obscure and counter-intuitive.

First, unlike in a plasma with equal numbers of \pm charges, gravity only attracts; no equivalent to Debye shielding limits interactions to roughly the inter-particle separation. Thus, interactions in a gravitational system are very long-range; indeed, the sum of weak ones from very distant stars dominates over nearby two-body encounters. Surrounded by $\gtrsim 10^5$ stars, each star sees a smooth potential and would need longer than the age of the Universe to exchange significant energy with others; self-gravitating systems are nearly collisionless. This implies that, if the potential has been stable since a galaxy formed, then a star retains memory of its original orbit and its observer can be a 'galactic archaeologist'. Conversely, if evidence suggests that certain orbit families have altered substantially, then by implication the galactic potential was disrupted significantly at least once in the past to scatter stars into new orbits.

Second, the stellar mean free path length exceeds greatly the system diameter, so its equation of state, pressure, and temperature cannot be defined. In contrast, charged plasma particles remain in a local equilibrium characterized by the ideal gas law.

As will soon be clear, solving the equations of a self-gravitating system is daunting. We touch only on key aspects, and refer the reader to the classic monograph of Binney and Tremaine (1987) to understand the full problem. We first introduce the distribution function

$f(\mathbf{x}, \mathbf{v}, t)$, the fine-grained probability of locating stellar mass in six-dimensional phase-space. Without collisions, the mass density within this piece of phase space is invariant $df/dt = 0$. Hence, the coupled Boltzmann and Poisson equations in gravitational potential Φ

$$\begin{aligned} \frac{df}{dt} &\equiv \frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla f - \nabla \Phi \cdot \frac{\partial f}{\partial \mathbf{v}} = 0 \\ \nabla^2 \Phi(\mathbf{x}, t) &= 4\pi G \int f(\mathbf{x}, \mathbf{v}, t) d^3 \mathbf{v} \end{aligned} \quad (\text{A.2})$$

describe the dynamics of collisionless, self-gravitating galaxies. Seven independent variables make f hard to solve. Instead, one often obtains insights from spatial or kinematical moments, yielding three partial differential equations. But, moments are *global* averages over the *entire* distribution, and there certainly *is* considerable material (both DM and possibly unsettled baryons) in the faint outer parts of galaxies, with unknown kinematics. In fact, this material is very important to understand because its slow rate of dynamical evolution (if undisturbed by tides) means that it may retain a memory of its initial state.

Appendix A.2.1. Velocity moments. Integrating the collisionless Boltzmann equation (CBE) (2) over velocity yields the Jeans equations

$$\frac{\partial \bar{v}_j}{\partial t} + \bar{v}_i \frac{\partial \bar{v}_j}{\partial x_i} = -\frac{\partial \Phi}{\partial x_j} - \frac{1}{\mathbf{v}} \frac{\partial (\mathbf{v} \sigma_{ij}^2)}{\partial x_i} \quad (\text{A.3})$$

where

$$\mathbf{v} \equiv \int f d^3 v \quad (\text{A.4})$$

is the mass density and

$$v_i \equiv \frac{1}{\mathbf{v}} \int f v_i d^3 v \quad (\text{A.5})$$

are the three velocity moments, and

$$-\mathbf{v} \sigma_{ij}^2 \equiv -\mathbf{v} (\bar{v}_i \bar{v}_j - \bar{v}_i \bar{v}_j) \quad (\text{A.6})$$

is the velocity dispersion stress tensor with

$$\bar{v}_i \bar{v}_j \equiv \frac{1}{\mathbf{v}} \int v_i v_j f d^3 v \quad (\text{A.7})$$

As mentioned, there is no equation of state to relate motions (through tensor σ^2) to mass density \mathbf{v} . Integrating higher moments $v_i v_k$ of the CBE over velocity does not make this link, it simply introduces an unspecified third-order tensor $\bar{v}_i \bar{v}_j \bar{v}_k$. The only route forward is to truncate the moment sequence, for example by making hopefully reasonable, but nonetheless non-unique, assumptions on σ^2 . This approach was first applied to galaxy formation by Larson (1969), who truncated and then integrated numerically the stellar and gas dynamical moment equations to model observed properties of Egs.

For example, a non-rotating galaxy with rotationally invariant densities and velocities has $\bar{v}_\theta^2(r) = \bar{v}_\phi^2(r)$. The anisotropy of tensor σ^2 is described by

$$\beta(r) = 1 - \bar{v}_\theta^2(r) / \bar{v}_r^2(r) \quad (\text{A.8})$$

with $\beta = -\infty, 0$, and 1 for orbits that are circular, isotropic, and radial, respectively. Even with this simplifying assumption, we unfortunately have only five equations to solve for six variables: \mathbf{v} , \bar{v}_r , \bar{v}_θ , Φ , \bar{v}_r^2 , and β (mass density, two independent mean velocities, gravitational potential, velocity dispersion along radii in the galaxy, and anisotropy of the

velocity dispersion). Next, we assume a spherically symmetrical Eg in a static configuration, i.e., without streaming motions. The Jeans equations then simplify to

$$\frac{1}{v} \frac{d(v\overline{v_r^2})}{dr} + 2 \frac{\beta \overline{v_r^2}}{r} = - \frac{d\Phi}{dr} \quad (\text{A.9})$$

Setting

$$\frac{\partial \Phi}{\partial r} = \frac{GM(r)}{r^2} \quad (\text{A.10})$$

recasts the radial Jeans equation into

$$v_c^2 = \frac{GM(r)}{r} = -\overline{v_r^2} \left(\frac{d \ln v}{d \ln r} + \frac{d \ln \overline{v_r^2}}{d \ln r} + 2\beta \right) \quad (\text{A.11})$$

to solve for $M(r)$ (see section 4.2 of Binney and Tremaine 1987; for details). If the radial variations of mass density, velocity dispersion, and velocity anisotropy β can be assumed or determined empirically by deprojecting measurements of luminous tracers such as planetary nebulae or globular clusters, we can recover the mass distribution. For example, Sargent *et al* (1978) assumed $\beta = 0$ (i.e., isotropic velocity dispersion) in the central regions of Eg M87, and found from the inward rise in both luminosity and velocity function dispersion that Υ must increase toward the centre, indicating a supermassive black-hole. Subsequently, Binney and Mamon (1982) relaxed the assumption $\beta = 0$, and found that rising Υ is unnecessary if large ad hoc variations in $\beta(r)$ are allowed.

Appendix A.2.2. Choice of velocity distribution function. Using only velocity moments does not ensure a viable solution of the Jeans equations (positive distribution function everywhere). Nor is any solution based on an assumed velocity dispersion tensor necessarily stable. An alternative approach to solving the CBE is to use prior knowledge of the distribution function. For a spherical galaxy, assume that the distribution function is both Maxwellian and ‘isothermal’, namely an isotropic velocity dispersion that is independent of radius. Unfortunately, such a galaxy has infinite mass and radius. When a dynamical system can be truncated tidally, a ‘lowered Maxwellian’ can be introduced. The distribution function of a lowered Maxwellian that truncates at positive energy ϵ is

$$f_K(\epsilon) = \begin{cases} \rho(2\pi\sigma^2)^{-3/2}(\exp[\epsilon/\sigma^2] - 1) & \epsilon > 0 \\ 0 & \epsilon \leq 0 \end{cases} \quad (\text{A.12})$$

Michie (1963) used (A.12) to model self-consistently the mass distributions of globular clusters, and King (1966) derived a model family that fit their SB profiles, although velocity function and proper motion measurements later showed anisotropic motions outside the cores. As discussed in section 3.2, observations show that Egs are actually triaxial, so have anisotropic velocity dispersions. Egs also have structured centres, not the constant density core of King models. Hence, King models are overly restrictive and are not fully realized in Egs.

Appendix A.2.3. Integrals of motion. The CBE can also be solved in terms of integrals of motion, namely independent functions of phase-space (\mathbf{x}, \mathbf{v}) that have constant value along any orbit. In a static potential (no dissipation), energy $E(\mathbf{x}, \mathbf{v}) = \frac{1}{2}v^2 + \Phi(\mathbf{x})$ is one and, if potential Φ is axisymmetric, the component of angular momentum around the rotation axis is too. In a spherical potential, angular momentum \mathbf{L} yields three integrals of motion. If the potential is central, there is a fifth integral whose five-dimensional surface in phase space $\psi(\mathbf{x}, \mathbf{v}) = \text{constant}$ intersects the 2D surface of constant E and \mathbf{L} to form a rosette-pattern orbit

that is restricted to a five-dimensional region of phase space. In a $1/r$ potential, the rosette closes and the fifth integral is said to be isolating. Otherwise, the entire 2D surface of constant E and \mathbf{L} is covered eventually by the non-isolating integral.

The dimensionality (hence complexity) of solutions to the CBE reduces greatly with solutions that depend only on isolating integrals. Orbits are regular if they have at least as many isolating integrals as spatial dimensions. Realistic potentials have fewer isolating integrals, hence have irregular orbits. The Strong Jeans Theorem states that a distribution function composed of regular orbits is a function of just three isolating integrals, $f(E, I_2, I_3)$.

For example, assuming that E is the only isolating integral in a spherical galaxy is equivalent to assuming an isotropic velocity dispersion tensor, which we noted is overly restrictive. To accord fully with the Strong Jeans Theorem requires a distribution function that depends on two further integrals without explicit forms. So, one next considers distribution functions that depend on both E and $|\mathbf{L}|$. Generally, the third integral is far harder to identify.

Appendix A.2.4. Stellar orbits and the gravitational potential. A galaxy can also be described by the orbit families of an assumed potential that reproduce its inferred distribution function and stellar motions.

An axisymmetric potential usually has a third integral with E and L_z . Some stars oscillate harmonically parallel to the cylindrical coordinate axes to form box orbits (Lissajous figures, which have no fixed sense of rotation and carry a star arbitrarily close to the centre of mass). At larger radii one finds loop orbits that are nearly circular and whose initial tangential velocity determines the elliptical annulus that confines the orbit. This resembles the annulus filled by an orbit in an axisymmetric potential as L_z is varied. Box orbits fill the phase space of more flattened axisymmetric systems. Orbital resonances, central density cusps, or central point masses change box orbits to boxlet orbits that tend to avoid small radii, or stochastic orbits that cover phase space chaotically when the central mass grows to $\sim 1\%$ of the system mass.

Realistic potentials and mass distributions thus have complex orbits to integrate numerically, but restricted potentials can be studied analytically. In particular, Stäckel (Eddington) potentials are the most general that separate in curvilinear coordinates, hence all orbits have three isolating integrals. Realistic galactic potentials (including those for triaxial figures) include both regular and irregular orbits, so Stäckel potentials are too restrictive. But their solutions are far more tractable because the moment equations now form a closed set. van de Ven *et al* (2003) discuss general solutions obtained with Stäckel potentials.

Appendix A.2.5. Spatial moments: tensor virial theorem. In Appendix A.2.1 we obtained three Jeans velocity moments by multiplying the CBE by each velocity component and then integrating over all velocities. Now multiply each velocity moment by a *position* coordinate and then integrate over all positions to form nine tensor virial theorem (tensor Virial Theorem) equations

$$\frac{1}{2} \frac{d^2 I_{jk}}{dt^2} = 2T_{jk} + \Pi_{jk} + W_{jk} \quad (\text{A.13})$$

with the moment of inertia tensor

$$I_{jk} = \int \rho x_j x_k d^3x \quad (\text{A.14})$$

partitioned into the potential energy tensor

$$W_{jk} = - \int \rho(x) x_j \frac{\partial \Phi}{\partial x_k} d^3x \quad (\text{A.15})$$

and the kinetic energy tensor

$$K_{jk} = \frac{1}{2} \int \rho \overline{v_j v_k} d^3x = T_{jk} + \frac{1}{2} \Pi_{jk} \quad (\text{A.16})$$

that sums ordered T and random Π motions, respectively. A ‘cold’ dynamical system is one where W_{jk} and T_{jk} balance, a ‘hot’ system balances W_{jk} and Π_{jk} . The tensor Virial Theorem links the global spatio-kinematical properties of a galaxy. While spatial data are straightforward to obtain in projection, kinematics were limited until recently to samples along a PA’s through narrow slits. Complete spatio-kinematical coverage over the inner part of an Eg is now attained in a single pointing of an integral-field spectrograph (IFS).

The most important insight from the tensor Virial Theorem follows when gravity collects matter from rest. As matter equilibrates, the tensor Virial Theorem tells us that half of the gravitational energy released causes motion within the potential well and the other half is dissipated to achieve binding energy $E_b = -E$ equal to the kinetic energy; the system has negative specific heat. For the MWg of mass M_g and average rotational velocity $\Theta_0 \approx 200 \text{ km s}^{-1}$. $E_b = K \approx \frac{1}{2} M_g \Theta_0^2$, hence $\approx 2.5 \times 10^{-7}$ of its rest mass energy must have dissipated during its formation.

The tensor Virial Theorem can estimate the global Υ of a nonrotating spherical galaxy by averaging stellar velocity dispersion $\sigma(R)$ and SB $\sigma(R)$ maps over the projected radius R from IFS spectra

$$\hat{I} = 3\pi \int_0^\infty \sigma(r) \sigma^2(r) r dr, \quad (\text{A.17})$$

and by recognising that density follows from an Abel inversion integral of the measured radial variation of the SB. With \bar{I} an integral of $\sigma(R)$, the tensor Virial Theorem then gives

$$M/L \equiv \Upsilon = -\frac{2\hat{I}}{\bar{I}} \quad (\text{A.18})$$

In section 3 the tensor Virial Theorem constrains Eg 3D shapes by comparing radial trends of mean rotational/random motion with SB ellipticity. tensor Virial Theorem volume averaged properties require non-unique assumptions.

Appendix A.2.6. Violent relaxation. Statistical mechanical quantities that depend on randomization such as system temperature and entropy are seemingly undefinable in a collisionless stellar system. Yet, virialization implies past dissipation. Violent relaxation Lynden-Bell (1967) is the dissipator that erases most initial conditions by widening the global range of stellar energies (but perhaps not the energies of certain orbit families), independent of stellar mass. It arises from rapid fluctuations in the gravitational field during hierarchical mass buildup that scatters the stellar energies. Simulations (for example Dekel *et al* 2005) show that the baryons approach virial equilibrium after a few dynamical times, namely a Maxwell distribution that is nearly isotropic near the centre

$$f(\epsilon) = \frac{\rho}{(2\pi\sigma^2)^{3/2}} \exp\left(-\frac{\Psi - v^2/2}{\sigma^2}\right), \quad (\text{A.19})$$

with isothermal equation of state $p(\rho) = K\rho$ and with Ψ the relative potential. Meanwhile, stars outside remain anisotropic, with many on radial orbits that originated in the mass buildup (see section 3.5).

In the CDM paradigm, DM relaxes to form the potential well. The baryons accrete into the well — their dissipation characterized by parameter

$$\lambda = \hat{J}|E|^{1/2} \frac{1}{GM^{3/2}} \quad (\text{A.20})$$

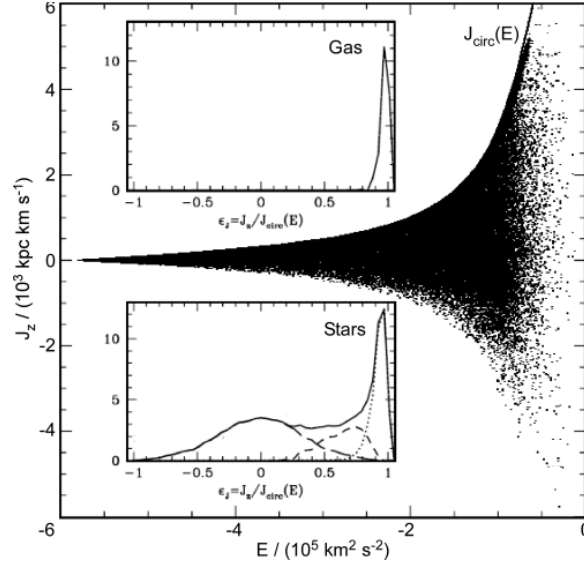


Figure A3. Abadi *et al* (2003)'s simulated formation of an Sa galaxy. The z-component of specific angular momentum within R_e of the baryonic galaxy is plotted versus the specific binding energy. Inserts plot the distribution of angular momentum (normalized to a fraction of circular orbit value) of gas and stars, together with a possible decomposition into different components (labelled \cdots for thin disc, $---$ for thick disc, $—$ for spheroid). Despite photometric resemblance to Sa galaxy components, the simulated rotation curve is too centrally peaked.

with $\hat{J} \equiv J/M$, which sets (for example Dalcanton *et al* 1997) the SB of the baryonic disc. Figure A3 shows the distribution of angular momentum from a Λ -CDM simulation (Abadi *et al* 2003), but present data are too sparse to constrain such models.

As discussed in section 4, \hat{J} of a disc spreads locally by collective motions from dynamical resonances in quasi-periodic orbits. Some DM distributions have minimal resonances hence maximal \hat{J} at the disc rim. We cannot yet measure stellar kinematics at this SB, but can see if the stellar disc appears to truncate (section 5.1.2).

Appendix A.2.7. Building a mass distribution and potential self-consistently from stellar orbits. To model self-consistently a galaxy, one combines orbit families of assumed Υ with a set of global weights to reproduce the distribution function as derived from the SB profile and the velocity function including asymmetries in absorption line profiles. Elucidating radial variations of Υ would be an important insight from a successful model, but current data are inadequate. At best, one posits a mass distribution with a spherical DM halo and stars whose average Υ is consistent with the observed radial profiles of absorption line strengths and photometric colours. The distribution function is assumed to be steady state, but recall that galaxy shape may arise from both initial conditions and external perturbations. To model an Eg one assumes inclination i and then

- use either a constraining technique such as maximum entropy, or, fit to a functional form that has an analytical inversion, to deproject the mapped SB $\sigma(R, z)$ with the Abel integral

$$\rho(r, z) = -\frac{\Upsilon}{\pi} \int_r^\infty \frac{d\sigma(R, z)}{dR} \frac{dR}{\sqrt{R^2 - r^2}} \quad (\text{A.21})$$

- calculate the corresponding potential $\Phi(R, z)$ assuming constant Υ ;
- solve Jeans equations for mean velocities, then divides random and streaming motions;
- integrate space motions to get the velocity function and dispersion;
- compare to data and iterate to refine Υ , the velocity division, and i . A valid model has positive distribution function throughout the galaxy phase space, while satisfying (Appendix A.2).

Velocity dispersions along the galaxy minor axis are always observed to be smaller than predicted by this procedure, implying a third integral and/or triaxial shape (section 3.2). Surface photometry finds non-elliptical isophotes for spheroidal galaxies. Some patterns are consistent with embedded discs, which may include dust and gas moving on inclined rings and counter-rotating compared to the stars. The cold ISM settles into stable orbits by differential (with radius) precession that collides clouds and dissipates their kinetic energy. The ISM may form a warped disc in an oblate potential or polar ring if prolate. An example is the dusty, but optically thin in near-IR, warped disc in Centaurus A (NGC 5128), studied recently with *SST* (Quillen *et al* 2006); based on transient tidal features (the warp and outer stellar shells), it may have arisen as a small gas-rich satellite merged ~ 0.1 Gyr ago.

Studies that assume a fixed, or slowly varying, potential from a mass model $\rho(r)$ use the orbit summing method of Schwarzschild (1979). Many orbits are computed for many oscillations in the specified potential; most are tubes or boxes without net rotation. By noting regions within an equipotential surface that some orbits never traverse, Schwarzschild found three integrals of motion for most orbits in his triaxial model potential. Next he summed orbits with positive weights to build the assumed $\rho(r)$. He found the weights by linear programming, but now maximum entropy and the Lucy (1974) or related stack-smooth-iterate (Bak and Statler 2000) algorithms are used. Methods such as ‘dithered’ initial conditions for orbit integration (Cappellari *et al* 2006) ensure the smooth distribution functions required to fit modern data; today over 400,000 orbits are summed with weights selected to reproduce the stellar density and velocity function.

Different distribution functions are consistent with a given mass model because the velocity function is degenerate with galaxy inclination unless additional features such as a stellar disc, ionized gas (for example Krajonović *et al* 2005), neutral gas (Morganti *et al* 2006), or dust lanes (Quillen *et al* 2006) also constrain inclination. If the mass distribution and potential are constrained by x-ray data, the orbital anisotropy is no longer degenerate. Today Jeans models are often used because of their transparency, uniqueness, and efficiency to generate an ensemble of noise-free distribution functions across a grid of structural parameters for Monte Carlo error analysis of the more sophisticated but non-unique Schwarzschild technique (for example Krajonović *et al* 2005).

Appendix B. A Brief Review of Hierarchical Galaxy Formation

Hierarchical buildup from small to large mass develops cosmic structure from the early linear phase explored in the spectrum of temperature fluctuations in the CMBR into the non-linear regime. In particular (for a comprehensive review, see Baugh 2006), Λ -CDM matches the power spectrum of the fluctuations as mapped by WMAP and terrestrial instruments, the Hubble diagram for high redshift SNe Ia, and the galaxy correlation function in recent wide-angle galaxy surveys. *Direct* study of underlying DM halos is impossible. Instead, we study the baryonic component that dissipates energy within the DM halo, forms stars, and leaves hot and cool gas. Key for a researcher interested only in the halo is to determine to what extent light traces DM. Appendix A introduces some of the galactic astrophysics that intrudes.

Galaxy and cluster build up in the Λ -CDM framework has been explored by simulations that follow up to 10^9 DM ‘particles’ from post-recombination density perturbations to the present epoch (for example Baugh 2006). Volumes ~ 500 Mpc on a side have been simulated with DM particles as small as $10^8 M_\odot$. Because DM is cold and non-interacting, gravitational N-body simulations can track its concentration into halos. In this way galaxy masses as small as $10^9 M_\odot$ are followed, unfortunately not the dwarf end of the galaxy mass function. Alternatively, galaxy build up can be followed in a statistical sense through semi-analytic prescriptions founded on the work of Press and Schechter (1974). The consensus in comparing both methods is that the predictions of Λ -CDM simulations are robust. However, study of baryonic matter within CDM halos requires accurate models of galaxy astrophysics over the mass range encompassed locally by a box ~ 12 Mpc on a side. For example, almost all SF seen at high redshifts comes from a cooling flow into CDM halos, yet its non-linear effects on the surrounding ISM (‘feedback’, see section 6.2) is well beyond present simulations.

A successful model must also reproduce the fate of baryons as galaxies merge. Helly *et al* (2003) compares the two tactics in use: 1) solve simultaneously the equations of hydrodynamics for gaseous baryons during the N-body simulation of the DM halo (for example Katz *et al* 1992, Steinmetz and Navarro 1999, Cen and Ostriker 2000, Pearce *et al* 2001). Naturally, at some point considerable sub-grid astrophysics must be parametrized. 2) graft this parametrized astrophysics onto statistical semi-analytic models (for example Cole, Baugh and Frenk 2000). We are still far from fundamental treatment of the gaseous baryons.

The starting point as the DM halo forms is to cool and settle baryonic gas into a SF disc. An Eg forms when two near-equal mass disc galaxies merge. It can later grow a new disc to form a large-bulge Sg only if hot halo gas remaining from the merger can cool and settle. The scenario reproduces the scale of early-type galaxies by forming galaxies in DM halos where baryon cooling is efficient: objects lighter than $10^{12} M_\odot$ form galaxies, those heavier form clusters. A persistent problem is that this scenario concentrates baryons too centrally, the so-called ‘angular momentum catastrophe’. DM halos acquire much of their angular momentum in mergers from the orbits of the progenitor galaxies. The baryonic gas, in cooling early toward the centres of the DM halos, does not gain this angular momentum (for example Steinmetz and Navarro 1999, Navarro and White 1994). Λ -CDM simulations therefore predict very structured halos and a central cusp of DM.

Despite these difficulties, simulations have provided robust results. Specifically, larger DM halos tend to assemble at lower redshifts (i.e., more recently) than smaller ones, and in general DM halos form more rapidly in denser than sparser environments (Baugh 2006).

The situation for baryons, particularly baryons now in stars, is murkier. Early simulations predicted that they would indeed follow DM through a process of adiabatic compression. More massive Egs are predicted to have younger luminosity-weighted mean ages than lower mass Egs, and those in clusters were predicted to be older in the mean than those in sparser regions. Egs of given mass are predicted to scatter significantly in luminosity-weighted mean age because of the stochastic nature of mergers (for example Kauffman and Charlot 1998). However, as we discuss in section 3 and section 4, the inferred mean ages of Egs indicate that low mass Egs typically contain younger stars than massive Egs. This anti-hierarchical behaviour (called ‘down-sizing’) led to new prescriptions for cooling and feedback (see section 6.2) related to SF and active galactic nuclei (AGN). For example, de Lucia *et al* (2005) suppress cooling of massive Egs that ends SF early on, while Croton *et al* (2006) and Bower *et al* (2006) use AGN for mass-dependent feedback to inhibit hierarchical mass build up of baryons.

In summary, observations are now connecting with models but both are still too uncertain to establish firmly the viability of the Λ -CDM scenario in the non-linear regime.

References

- Abadi M G, Navarro J F, Steinmetz M and Eke V R 2003 *Astrophys. J.* **597** 21–34
- Abadi M G, Moore B and Bower R G 1999 *Mon. Not. Roy. Ast. Soc.* **308** 947–54
- Alcock C *et al* 2000 *Astrophys. J.* **542** 281–307
- Antia H M and Basu S 2006 *Astrophys. J.* **644** 1292–8
- Ashman K E and Zepf S E 1992 *Astrophys. J.* **384** 50–61
- Asplund M, Grevesse N, Sauval A J, Allende Prieto C and Kiselman D 2005 *Astron. Astrophys.* 435 339–40
- Athanassoula E 2003 *Mon. Not. Roy. Ast. Soc.* **341** 1179–98
- Athanassoula E and Bosma A 2003 *Astrophys. Space Sci.* **284** 491–4
- Aubourg E *et al* 1993 *Nature* **365** 623–5
- de Avillez M A 2000 *Mon. Not. Roy. Ast. Soc.* **315** 479–97
- Azzaro M *et al* 2007 *Mon. Not. Roy. Ast. Soc.* **376** L43–7
- Babusiaux C and Gilmore G 2005 *Mon. Not. Roy. Ast. Soc.* **358** 1309–19
- Bahcall J N, Basu S and Serenelli A M 2005 *Astrophys. J.* **631** 1281–5
- Bak J and Statler T S 2000 *Astron. J.* **120** 110–22
- Baldry I K, Balogh M L, Bower R, Glazebrook K and Nichol R C 2004 *AIP Conf. Proc.* vol 743 (Melville, NY: AIP) p 106–19 (*Preprint astro-ph/0410603*)
- Balogh M L, Morris S L, Yee H K C, Carlberg R G and Ellingson E 1999 *Astrophys. J.* **527** 54–79
- Barbuy B 1999 *Astrophys. Space Sci.* **265** 319–26
- Bartelmann M and Schneider P 2001 *Phys Reports* **340** 291
- Barth A J 2007 *Astron. J.* **133** 1085–91
- Baugh C M 2006 *Rep. Prog. Phys.* **69** 3101–56
- Bautz M W and Arabadjis J S 2003 *Clusters of Galaxies: Probes of Cosmological Structure and Galaxy Evolution* (*Carnegie Observatories Astrophysics Series* vol 3) ed J S Mulchaey, A Dressler and A Oemler (Cambridge: Cambridge Univ. Press) (*Preprint astro-ph/0303313*)
- Beaton R L *et al* 2007 *Astrophys. J.* **658** L91–4
- Bekki K and Chiba M 2000 *Astrophys. J.* **534** L89–92
- Bekki K, Koribalski B S and Kilborn V A 2005 *Mon. Not. Roy. Ast. Soc.* **363** L21–5
- Bell E F and de Jong R S 2001 *Astrophys. J.* **550** 212–29
- Belokurov V *et al* 2006 *Astrophys. J.* **642** L137–40
- Bender R *et al* 2005 *Astrophys. J.* **631** 280–300
- Bennett D P, Becker A C and Tomaney A 2005 *Astrophys. J.* **631** 301–11
- Benson A J, Bower R G, Frenk C S, Lacey C G, Baugh C M and Cole S 2003 *Astrophys. J.* **599** 38–49
- Berentzen I, Shlosman I, Martinez-Valpuesta I and Heller C 2007 (*Preprint astro-ph/0703028*)
- Bertin G and Lin C C 1996 *Spiral Structure in Galaxies* (Cambridge, MA: MIT Press)
- Bertola F and Capaccioli M 1975 *Astrophys. J.* **200** 439–45
- Bessell M S, Castelli F and Plez B 1998 *Astron. Astrophys.* **333** 231–50
- Bicknell G, Sutherland R S, van Breugel, W J M, Dopita M A, Dey A and Miley G K 2000 *Astrophys. J.* **540** 678–86
- Binggeli B, Sandage A and Tammann G A 1985 *Astron. J.* **90** 1681–759
- Binney J 1978 *Mon. Not. Roy. Ast. Soc.* **183** 501–14
- Binney J J and Mamon G A 1982 *Mon. Not. Roy. Ast. Soc.* **200** 361–75
- Binney J and Tremaine S 1987 *Gravitational Dynamics* (Princeton: Princeton University Press)
- Binney J and Merrifield M 1998 *Galactic Astronomy* (Princeton: Princeton University Press)
- Binney J 2005 *Mon. Not. Roy. Ast. Soc.* **363** 937–42
- Bissantz N and Gerhard O 2002 *Mon. Not. Roy. Ast. Soc.* **330** 591–608
- Blackwell D E and Lynas-Gray A E 1998 *Astron. Astrophys.* **129** S505–15
- Bland-Hawthorn J, Freeman K C and Quinn P J 1997 *Astrophys. J.* **490** 143–55
- Bland-Hawthorn J and Freeman K C 2004 *The Fifth Workshop on Galactic Chemodynamics Publ. Astro. Soc. Australia* **21** pp 110–20
- Bland-Hawthorn J, Vlaré M, Freeman K C and Draine B T 2005 *Astrophys. J.* **629** 239–49
- Blitz L, Fukui Y, Kawamura A, Leroy A, Mizuno N and Rosolowsky E 2006 *Preprint astro-ph/0602600*
- Borkova T V and Marsakov V A 2003 *Astron. Astrophys.* **398** 133–9
- Bosma A 1981 *Astron. J.* **86** 1825–46
- Bothun G, Impey C and McGaugh S 1997 *Publ. Astron. Soc. Pac.* **109** 745–58
- Bottema R 1993 *Astron. Astrophys.* **275** 16–36
- Bournaud F, Combes F and Semelin B 2005 *Mon. Not. Roy. Ast. Soc.* **364** L18–22
- Bower R G, Benson A J, Malbon R, Helly J C, Frenk C S, Baugh C M, Cole S and Lacey C G 2006 *Mon. Not. Roy. Ast. Soc.* **370** 645–55
- Brinchmann J, Charlot S, White S D M, Tremonti C, Kauffmann G, Heckman T and Brinkmann J 2004 *Mon. Not. Roy. Ast. Soc.* **351** 1151–79

- Broeils A H 1992 *Astron. Astrophys.* **256** 19–32
- Brown T M, Smith E, Guhathakurta P, Rich R M, Ferguson H C, Renzini A, Sweigart A V and Kimble R A 2006 *Astrophys. J.* **636** L89–92
- Bruzual G and Charlot S 2003 *Mon. Not. Roy. Ast. Soc.* **344** 1000–34
- Burstein D 1979 *Astrophys. J.* **234** 829–36
- Buta R and Combes F 1996 *Fundamentals of Cosmic Physics* **17** 95–281
- Butcher H and Oemler A 1978 *Astrophys. J.* **226** 559–65
- Caldwell N 1983 *Astron. J.* **88** 804–12
- Caldwell N, Rose J A, Sharples R M, Ellis R S and Bower R G 1993 *Astron. J.* **106** 473–92
- Caldwell N C and Rose J A 1997 *Astron. J.* 492–520
- Caldwell N, Rose J A and Dendy K 1999 *Astron. J.* **117** 140–56
- Caldwell N, Rose J A and Concannon K 2003 *Astron. J.* **125** 2891–926
- Capaccioli M, Held E V, Lorenz H and Vietri M 1990 *Astron. J.* **99** 1813–22
- Capaccioli M, Vietri M, Held E V and Lorenz H 1991 *Astrophys. J.* **371** 535–40
- Cappellari M *et al* 2005 *Preprint astro-ph/0509470*
- Cappellari M *et al* 2006 *Mon. Not. Roy. Ast. Soc.* **366** 1126–50
- Cardiel N, Gorgas J, Sanchez-Blazquez P, Cenarro A J, Pedraz S, Bruzual G and Klement J 2003 *Astron. Astrophys.* **409** 511–22
- Carignan C, Chemin L, Huchtmeier W K and Lockman F J 2006 *Astrophys. J.* **641** L109–12
- Carney B W, Laird J B, Latham D W and Aguilar L A 1996 *Astron. J.* **112** 668–92
- Carrera R, Aparicio A, Martinez-Delgado D and Alonso-Garcia J 2002 *Astron. J.* **123** 3199–209
- Carollo C M 2005 *Co-evolution of black holes and galaxies (Carnegie Observatory Astrophysics Series vol 1)* ed L C Ho (Cambridge: Cambridge University Press) pp 231–47
- Cattaneo A, Blaizot J, Weinberg D H, Colombi S, Dave R, Devriendt J, Guiderdoni B, Katz N and Keres D 2007 *Mon. Not. Roy. Ast. Soc.* **377** 63–76
- Cattaneo A, Dekel A, Devriendt J, Guiderdoni B and Blaizot J 2006 *Mon. Not. Roy. Ast. Soc.* **370** 1651–65
- Cecil G., Bland-Hawthorn J, Veilleux S and Filippenko A V 2001 *Astrophys. J.* **555** 338–55
- Cen R and Ostriker J P 2000 *Astrophys. J.* **538** 83–91
- Cerviño M, Mas-Hesse J M and Kunth D 2002 *Astron. Astrophys.* **392** 19–31
- Chapman S, Ibata R, Lewis G F, Ferguson A M N, Irwin N, McConnachie A and Tavnir N 2006 *Astrophys. J.* **653** 255–66
- Chiba T and Beers T 2001 *Astrophys. J.* **549** 325–36
- Clowe D, Bradac M, Gonzalez A H, Markevitch M, Randall S W, Jones C and Zaritsky D 2006 *Astrophys. J.* **648** L109–11
- Coelho P, Barbuy B, Melendez J, Schiabon R P and Castilho B 2005 *Astron. Astrophys.* **443** 735–46
- Cole S, Lacey C G, Baugh C M and Frenk C S 2000 *Mon. Not. Roy. Ast. Soc.* **319** 168–204
- Colless M *et al* 2001 *Mon. Not. Roy. Ast. Soc.* **328** 1039–63
- Cooper M C *et al* 2006 *Mon. Not. Roy. Ast. Soc.* **370** 198–212
- Courteau S, de jong R S and Broeils A H 1996 *Astrophys. J.* **457** L73–6
- Courteau S and Rix H 1999 *Astrophys. J.* **513** 561–71
- Courteau S, Andersen D R, Bershadsky M A, MacArthur L A and Rix H-W 2003 *Astrophys. J.* **594** 208–24
- Côté S, Carignan C and Freeman K C 2000 *Astron. J.* **120** 3027–59
- Cox T J, Primack J, Jonsson P and Somerville R S 2004 *Astrophys. J.* **607** L87–90
- Cox D P 2005 *Ann. Rev. Astron. Astrophys.* **43** 337–86
- Crézé M, Chereul E, Bienayme O and Pichon C 1998 *Astron. Astrophys.* **329** 920–36
- Croft S *et al* 2006 *Astrophys. J.* **647** 1040–55
- Croton D J, Springel V, White S D M, de Lucia G, Frenk C S, Gao L, Jenkins A, Kauffmann G, Navarro J F and Yoshida N 2006 *Mon. Not. Roy. Ast. Soc.* **365** 11–28
- Crowl H H, Kenney J D P, van Gorkom J H and Vollmer B 2005 *Astrophys. J.* **130** 65–72
- Dalcanton J J, Spergel D N and Summers F J 1997 *Astrophys. J.* **482** 659–76
- Davies R L, Efstathiou G, Fall S M, Illingworth G and Schechter P L 1983 *Astrophys. J.* **266** 41–57
- Davies R L and Birkinshaw M 1988 *Astrophys. J. Supp.* **68** 409–47
- de Blok W J G and McGaugh S S 1996 *Astrophys. J.* **469** L89–92
- de Blok W J G and McGaugh S S 1997 *Mon. Not. Roy. Ast. Soc.* **290** 533–52
- de Grijs R, Kregel M and Wesson K H 2001 *Mon. Not. Roy. Ast. Soc.* **324** 1074–86
- de Lucia G, Springel V, White S D M, Croton D and Kauffmann G 2006 *Mon. Not. Roy. Ast. Soc.* **366** 499–509
- de Silva G M *et al* 2006 *Astron. J.* **131** 455–60
- de Vaucouleurs G 1948 *Ann. Astrophys.* **11** 247–87
- Debatista V P, Mayer L, Carollo C M, Moore B, Wadsley J and Quinn T 2006 *Astrophys. J.* **645** 209–27
- Dekel A, Stoehr F, Mamon G A, Cox T J, Novak G S and Primack J R 2005 *Nature* **437** 707–10
- Di Benedetto G P 1998 *Astron. Astrophys.* **339** 858–71

- Dolphin A E 2002 *Mon. Not. Roy. Ast. Soc.* **332** 91–108
- Dominguez I, Chieffi A, Limongi M and Straniero O 1999 *Astrophys. J.* **524** 226–41
- Dopita M A, Groves B, Sutherland R S, Binette L and Cecil G 2002 *Astrophys. J.* **572** 753–61
- Dopita M A 2005 *The Spectral Energy Distribution of Gas-Rich Galaxies: Confronting Models with Data* ed Popescu C C and Tuffs R J (Melville, NY: AIP) *AIP Conf. Proc.* vol 761 p 203–22 (Preprint astro-ph/0502339)
- Dopita M A, Fischera A, Sutherland R S, Kewley L J, Tuffs R J, Popescu C C, van Breugel W, Groves B A and Leitherer C 2006 *Astrophys. J.* **647** 244–55
- Douglas N G *et al* 2002 *Publ. Astron. Soc. Pac.* **114** 1234–51
- Dressler A 1980 *Astrophys. J.* **236** 351–65
- Dressler A 1986 *Astrophys. J.* **301** 35–43
- Dressler A, Oemler A Jr, Couch W J, Smail I, Ellis R S, Barger A, Butcher H, Poggianti B M and Sharples R M 1997 *Astrophys. J.* **490** 577–91
- Dressler A, Oemler A Jr, Poggianti B M, Smail I, Trager S, Shectman S A, Couch W J and Ellis R S 2004 *Astrophys. J.* **617** 867–78
- Driver S P, Popescu C, Tuffs R J, Liske J, Graham A, De Propriis R and Allen P D 2007 *Mon. Not. Roy. Ast. Soc.* in press
- Dwek E, Arendt R G, Hauser M G, Kelsall T, Lisse C M, Moseley S H, Silverberg R F, Sodroski T J and Weiland J L 1995 *Astrophys. J.* **445** 716–30
- Eggen O J, Lynden-Bell D and Sandage A R 1962 *Astrophys. J.* **136** 748–66
- Eisenhauer F, Schoedel R, Genzel R, Ott T, Tecza M, Abuter R, Eckart A and Alexander A 2003 *Astrophys. J.* **597** L121–4
- Elmegreen B and Falgarone E 1996 *Astrophys. J.* **471** 816–21
- Epchtein N *et al* 1994 *Astr. Sp. Sci.* **217** 3–9
- Erwin P, Beckman J E and Pohlen M 2005 *Astrophys. J.* **626** L81–4
- Eskridge P B *et al* 2000 *Astron. J.* **119** 536–44
- Evans N J 1999 *Ann. Rev. Astron. Astrophys.* **37** 311–62
- Evans N W and Belokurov V 2002 *Astrophys. J.* **567** :119–23
- Faber S M 1973 *Astrophys. J.* **179** 731–54
- Faber S M, Friel E D, Burstein D and Gaskell C M 1985 *Astrophys. J. Supp.* **57** 711–41
- Faber S M *et al* 2004 (Preprint) astro-ph/0506044
- Fabian A C 1994 *Ann. Rev. Astron. Astrophys.* **32** 277–318
- Fellhauer M *et al* 2006 *Astrophys. J.* **651** 167–73
- Feltzing S, Bensby T and Lundstrom I 2003 *Astron. Astrophys.* **397** L1–4
- Ferguson H C 1989 *Astron. J.* **98** 367–418
- Ferrarese L and Merritt D 2000 *Astrophys. J.* **539** L9–12
- Ferreras I, Pasquali A, de Carvalho R R, de la Rosa I G and Lahav O 2006 *Mon. Not. Roy. Ast. Soc.* **370** 828–36
- Ferreras I, Saha P and Williams L I R 2005 *Astrophys. J.* **623** L5–8
- Ferrière K M 2001 *Rev. Mod. Phys.* **73** 1031–66
- Font A, Johnston K, Guhathakurta R, Majewski S and Rich M 2006 *Astron. J.* , **131** 1436
- Fragile P C, Murray S D, Anninos P and van Breugel W 2004 *Astrophys. J.* **604** 74–87
- Fraternali F and Binney J J 2006 *Mon. Not. Roy. Ast. Soc.* **366** 449–66
- Freeman K and Bland-Hawthorn J 2002 *Ann. Rev. Astron. Astrophys.* **40** 487–537
- Freeman K and McNamara G 2006 *In Search of Dark Matter* (Chichester: Praxis Publishing)
- Friedli D, Benz W and Kennicutt R 1994 *Astrophys. J.* **430** L105–8
- Fuhrmann K, Pfeiffer M, Frank C, Reetz J and Gehren T 1997 *Astron. Astrophys.* **323** 909–22
- Gallart C, Freedman W L, Aparicio A, Bertelli G and Chiosi C 1999 *Astron. J.* **118** 2245–61
- Gallart C, Zoccali M and Aparicio A 2005 *Ann. Rev. Astron. Astrophys.* **43** 387–434
- García-Ruiz I, Sancisi R and Kuijken K 2002 *Astron. Astrophys.* **394** 769–89
- Gavazzi G 1987 *Astrophys. J.* **320** 96–121
- Gavazzi G, Contursi A, Carrasco L, Boselli A, Kennicutt R, Scodreggio M and Jaffe W 1995 *Astron. Astrophys.* **304** 325–340
- Gebhardt K *et al* 2000 *Astrophys. J.* **539** L13–6
- Geha M, Guhathakurta P and van der Marel R P 2002 *Astron. J.* **124** 3073–87
- Ghez A M, Salim S, Hornstein S D, Tanner A, Lu J R, Morris M, Becklin E E and Duchêne G 2005 *Astrophys. J.* **620** 744–57
- Gilmore G and Reid N 1983 *Mon. Not. Roy. Ast. Soc.* **202** 1025–47
- Giovanelli R and Haynes M P 1985 *Astrophys. J.* **292** 404–25
- Glazebrook K and Bland-Hawthorn J 2001 *Publ. Astron. Soc. Pac.* **113** 197–214
- González-Delgado R M *et al* 1999 *Astrophys. J. Supp.* **125** 489–509
- Gordon M A and Sorooshenko R L 2003 *Radio Recombination Lines, Their Physics and Astronomical Applications* (New York: Springer)

- Gould A, Bahcall J N and Flynn C 1997 *Astrophys. J.* **482** 913–8
- Gould A, Flynn C and Bahcall J N 1998 *Astrophys. J.* **503** 798–808
- Graham A W and Guzmán R 2003 *Astron. J.* **125** 2936–50
- Graham A W *Near-fields cosmology with dwarf elliptical galaxies* (Cambridge: Cambridge U. Press) IAU Colloquium Proc. 198 pp 303–10
- Gratton R G, Carretta E, Desidera S, Lucatello S, Mazzei P and Barbieri M 2003 *Astron. Astrophys.* **406** 131–40
- Grebel, Gallagher and Harbeck 2003 *Astron. J.* **125** 1926–39
- Grebel E K 2005 *AIP Conf. Proc.* vol 752 (Melville, NY: AIP) p 161–74
- Guhathakurta R, Rich M R, Reitzel D B, Cooper M C, Gilbert K M, Majewski S R, Ostheimer J C, Geha M C, Johnston K V and Patterson R J 2006 *Astron. J.* **131** 2497–513
- Gunn J E and Gott R 1972 *Astrophys. J.* **176** 1–19
- Gunn J E, Stryker L L and Tinsley B M 1981 *Astrophys. J.* **249** 48–67
- Hafner L M, Reynolds R J, Tufte S L, Madsen G J, Jaehnig K P and Percival J W 2003 *Astrophys. J. Supp.* **149** 405–22
- Häring N and Rix H -W 2004 *Astrophys. J.* **604** L89–92
- Hartwick F D A 2002 *Astrophys. J.* **576** L29–32
- Heckman T M *et al* 2005 *Astrophys. J.* **619** L35–8
- Helly J C, Cole S, Frenk C S, Baugh C M, Benson A, Lacey C and Pearce F R 2003 *Mon. Not. Roy. Ast. Soc.* **338** 913–925
- Helmi A, White S D M and Springel V 2003 *Mon. Not. Roy. Ast. Soc.* **339** 834–48
- Henry T J, Ianna P A, Kirkpatrick J D and Jahreiss H 1997 *Astron. J.* **114** 388–95
- Ho L C, Filippenko A V and Sargent W L W 1997 *Astrophys. J.* **487** 568–78
- Hoekstra H, Yee H K C and Gladders M D 2004 *Astrophys. J.* **606** 67–77
- Hunt L K, and Malkan M A 1999 *Astrophys. J.* **516** 660–671
- Hunter C and Toomre A 1969 *Astrophys. J.* **155** 747–76
- Ibata K, Irwin M, Lewis G, Ferguson A and Tanvir N 2001 *Astrophys. J.* **551** 294–311
- Ibata K, Irwin M, Lewis G, Ferguson A and Tanvir N 2001 *Nature* **412** 49–52
- Ibata K, Irwin M J, Lewis G F, Ferguson A M N and Tanvir N 2003 *Mon. Not. Roy. Ast. Soc.* **340** L21–7
- Ibata K, Chapman S, Ferguson A, Irwin M and Lewis G 2004 *Mon. Not. Roy. Ast. Soc.* **351** 117–24
- Ibata K, Martin N F, Irwin M, Chapman S, Ferguson A M N, Lewis G F and McConnachie A W 2007 *Preprint astro-ph/0704.1318*
- Iben I 1974 *Ann. Rev. Astron. Astrophys.* **12** 215–56
- Iben I 1991 *Astrophys. J. Supp.* **76** 55–114
- Iglesias C A and Rogers F J 1996 *Astrophys. J.* **464** 943–53
- Illingworth G 1977 *Astrophys. J.* **218** L43–7
- Impey C and Bothun G 1997 *Ann. Rev. Astron. Astrophys.* **35** 267–307
- Ivezic Z, Lupton R H, Schlegel D, Jonston D, Gunn J E, Knapp G R, Strauss M A and Rockosi C M 2004 *Satellites and Tidal Streams* (San Francisco: Astro. Soc. Pacific) *ASP Conf. Proc.* vol 327 p 104
- Jetzer Ph and Novati S C 2004 *Preprint astro-ph/0407209*
- Jimenez R, Panter B, Heavens A F and Verde L 2005 *Mon. Not. Roy. Ast. Soc.* **356** 495–501
- Jones L, Smail I and Couch W 2000 *Astrophys. J.* **528** 118–22
- Jørgensen I, Franx M and Kjaergaard P 1996 *Mon. Not. Roy. Ast. Soc.* **280** 167–85
- Kallivayalil N, van der Marel R P and Alcock C 2006 *Astrophys. J.* **652** 1213
- Kallivayalil N *et al* 2006 *Astrophys. J.* **648** 772
- Kannappan S J, Jansen R A and Barton E J 2004 *Astron. J.* **127** 1371–85
- Katz N, Hernquist L and Weinberg D H 1992 *Astrophys. J.* **399** L109–12
- Kauffman G and Charlot S 1998 *Mon. Not. Roy. Ast. Soc.* **294** 705–17
- Kauffman G *et al* 2003 *Mon. Not. Roy. Ast. Soc.* **341** 33–53
- Kenney J and Koopmann 1999 *Astron. J.* **117** 181
- Kennicutt R C, Tamblyn, & Congdon 1994 *Astrophys. J.* **435** 22–36
- Kennicutt R C 1998 *Ann. Rev. Astron. Astrophys.* **36** 189–231
- Kent S M 1986 *Astron. J.* **91** 1301–27
- Keres D, Kaatz N, Weinberg D H and Davé R 2005 *Mon. Not. Roy. Ast. Soc.* **363** 2–28
- King I R 1962 *Astron. J.* **67** 471–85
- King I R 1966 *Astron. J.* **71** 64–75
- Kinman T D, Cacciari C, Bragaglia A, Buzzoni A and Spagna A 2007 *Mon. Not. Roy. Ast. Soc.* **375** 1381–98
- Kleyna J T, Wilkinson M I, Evans N W and Gilmore G 2001 *Astrophys. J.* **563** L115–8
- Kleyna J T, Wilkinson M I, Evans N W and Gilmore G 2005 *Astrophys. J.* **630** L141–4
- Knapen, J H 1999 *The Evolution of Galaxies on Cosmological Timescales* (San Francisco: Astro. Soc. Pacific) *ASP Conf. Proc.* vol 187 72–87
- Knapen, J H, Shlosman, I, and Peletier, R F 2000 *Astrophys. J.* **529** 93–100

- Kobayashi C and Arimoto N 1999 *Astrophys. J.* **527** 573–99
- Koch *et al* 2007 *Astron. J.* **133** 270–83
- Koch A and Grebel E K 2006 *Astron. J.* **131** 1405–15
- Koopmann R A and Kenney J D P 1998 *Astrophys. J.* **497** L75–9
- Kormendy J 1982 *Morphology and Dynamics of Galaxies, Proc. of the Twelfth Advanced Course Saas-Fee Switzerland Observatoire de Geneve* pp 113–288
- Kormendy J 1985 *Astrophys. J.* **295** 73–9
- Kormendy J and Richstone D 1995 *Ann. Rev. Astron. Astrophys.* **33** 581–624
- Kormendy J and Kennicutt R 2004 *Ann. Rev. Astron. Astrophys.* **42** 603–683
- Kormendy J and Fisher D B 2005, *Rev. Mex. A. A.* **23** 101–8
- Kormendy J, Cornell M E, Block D L, Knapen J H and Allard E L 2006 *Astrophys. J.* **642** 765–74
- Kozhurina-Platais V, Demarque P, Platais I, Orosz J A and Barnes S 1997 *Astron. J.* **113** 1045–56
- Krajonović D, Cappelari M, Emsellem E, McDermid R M and de Zeeuw P T 2005 *Mon. Not. Roy. Ast. Soc.* **357** 1113–33
- Krajonović D, Sharp R and Thatte N 2007 *Mon. Not. Roy. Ast. Soc.* **374** 385–98
- Kranz T, Slyz A and Rix H-W 2003 *Astrophys. J.* **586** 143–51
- Kroupa P 2001 *Mon. Not. Roy. Ast. Soc.* **322** 231–46
- Kroupa P 2002 *Mon. Not. Roy. Ast. Soc.* **330** 707–18
- Kuijken K and Rich R M 2002 *Astron. J.* **124** 2054–66
- Kuntschner H 2000 *Mon. Not. Roy. Ast. Soc.* **315** 184–208
- Lambas D G, Maddox S J and Loveday J 1992 *Mon. Not. Roy. Ast. Soc.* **258** 404–14
- Larson R B 1969 *Mon. Not. Roy. Ast. Soc.* **145** 405–22
- Larson R B and Tinsley B M 1978 *Astrophys. J.* **219** 46–59
- Larson R B, Tinsley B M and Caldwell C N 1980 *Astrophys. J.* **237** 692–707
- Laurikainen, E, Salo, H, and Buta, R 2004 *Astrophys. J.* **607** 103–124
- Lavery R J and Henry J P 1988 *Astrophys. J.* **330** 596–600
- Lebreton Y 2001 *Ann. Rev. Astron. Astrophys.* **38** 35–77
- Lee J-W and Carney B W 1999 *Astron. J.* **118** 1373–89
- Lee H C, Worthey G, Trager S C and Faber S M 2007 *Astrophys. J.* in press (*Preprint astro-ph/0605425*)
- Leitherer C, Schaerer D, Goldader J D, Delgado R M, González R C, Kune D F, de Mello D F, Devost D and Heckman T M 1999 *Astrophys. J. Supp.* **123** 3–40
- Levine E S, Blitz L and Heiles, C 2006 *Science* **312** 1773–7
- Levy L, Rose J A, van Gorkom J and Chaboyer B 2007 *Astron. J.* **133** 1104–24
- Lo K Y 2005 *Ann. Rev. Astron. Astrophys.* **43** 625–76
- Lockman F J 2002 *Seeing Through the Dust* (San Francisco: Astro. Soc. Pacific) *ASP Conf. Proc.* vol 276 p 107
- Lockman F J 2003 *Astrophys. J.* **591** L33–36
- Lombardi M *et al* 2005 *Astrophys. J.* **623** 42–56
- Lucy L R 1974 *Astron. J.* **79** 745–54
- Lynden-Bell D 1967 *Mon. Not. Roy. Ast. Soc.* **136** 101–21
- Lynden-Bell D and Wood R 1968 *Mon. Not. Roy. Ast. Soc.* **138** 495–525
- Lynden-Bell D and Kalnajs A J 1972 *Mon. Not. Roy. Ast. Soc.* **157** 1–30
- McConnachie A W, Chapman S, Ibata R A, Ferguson A M N, Irwin M K, Lewis G F and Tanvir N R 2006 *Astrophys. J.* **647** L25–8
- McNamara B R, Nulsen P E J, Wise M W, Rafferty R A, Carilli C, Sarazin C L and Blanton E L 2005 *Nature* **433** 45–7
- MacArthur L A, Courteau S and Holtzman J A 2003 *Astrophys. J.* **582** 689–722
- MacArthur L, Courteau S, Bell E and Holtzman J A 2004 *Astrophys. J. Supp.* **152** 175–99
- Maccarone T J, Fender R P and Tzioumis A K 2004 *Preprint astro-ph/0412014*
- Madgwick D S, Somerville R, Lahav O, and Ellis R 2003 *Mon. Not. Roy. Ast. Soc.* **343** 871–9
- Majewski S R, Munn J A and Hawley S L 1996 *Astrophys. J.* **459** L73–7
- Majewski S R, Skrutskie M F, Weinberg M D and Ostheimer J C 2003 *Astrophys. J.* **599** 1082–1115
- Majewski S R 2004 *Astron. J.* **128** 245–59
- Malhotra S 1995 *Astrophys. J.* **448** 138–48
- Maraston C, Greggio L, Renzini A, Ortolani S, Saglia R P, Puzia T H and Kissler-Patig M 2003 *Astron. Astrophys.* **400** 823–40
- Marchant A B and Olson D W 1979 *Astrophys. J.* **230** L157–9
- Marcolini A, Strickland D K, D’Ercole A, Heckman T M and Hoopes C G 2005 *Mon. Not. Roy. Ast. Soc.* **362** 626–48
- Marconi A and Hunt L K 2003 *Astrophys. J.* **589** L21–4
- Marinova, I, and Jogee, S 2007 *Astrophys. J.* **659** 1176–1197
- Markevitch M, Gonzalez, A H, Viklenin A, Murray S, Forman W, Jones C and Tucker W 2002 *Astrophys. J.* **567** L27–31

- Martins L P, Gonzalez Delgado R M, Leitherer C, Cervino M and Hauschildt P 2005 *Mon. Not. Roy. Ast. Soc.* **358** 49–65
- Mateo M 1998 *Ann. Rev. Astron. Astrophys.* **36** 435–506
- Mathewson D S, Ford V L and Buchhorn M 1992 *Astrophys. J. Supp.* **81** 413–659
- McGaugh S, Rubin V C and de Blok W J G 2001 *Astron. J.* **122** 2381–95
- McGaugh S M 2004 *Astrophys. J.* **609** 652–66
- McGaugh S S 2005 *Astrophys. J.* **632** 859–71
- McWilliam A and Rich R M 1994 *Astrophys. J. Supp.* **91** 749–91
- McWilliam A 1997 *Ann. Rev. Astron. Astrophys.* **35** 503–56
- Merritt D 1993 *Astrophys. J.* **413** 79–94
- Merritt D 2006 *Rep. Prog. Phys.* **69** 2513–79
- Mihos J C and Hernquist L 1996 *Astrophys. J.* **464** 641–63
- Mihos J C, Spaans M and McGaugh S S 1999 *Astrophys. J.* **515** 89–96
- Minchin R *et al* 2005 *Astrophys. J.* **622** L21–4
- Michie R W 1963 *Mon. Not. Roy. Ast. Soc.* **125** 127–39
- Momany Y, Zaggia S R, Bonifacio P, Piotto G, De Angeli F, Bedin L R and Carraro G. 2004 *Astron. Astrophys.* **421** L29–32
- Monelli M *et al* 2003 *Astron. J.* **126** 218–36
- Moore B, Katz N, Lake G, Dressler A and Oemler A 1996 *Nature* **379** 613–6
- Moorthy B K and Holtzman J A 2006 *Mon. Not. Roy. Ast. Soc.* **371** 583–608
- Morganti R, de Zeeuw P T, Oosterloo T A, McDermid R M, Krajnović D, Cappellari M, Kenn F and Weijmans A 2006 *Mon. Not. Roy. Ast. Soc.* **371** 157–69
- Moultaka J, Ilovaisky S A, Prugniel P and Soubiran C 2004 *Publ. Astron. Soc. Pac.* **116** 693–8
- Mouawad N, Eckart A, Pfalzner S, Schödel R, Moultaka J and Spurzem R 2005 *Astronomische Nachrichten* **326** 83–95
- Muchale J 2000 *Ann. Rev. Astron. Astrophys.* **38** 289–335
- Mulchaey, J S, and Regan, M W 1997 *Astrophys. J.* **482** L135–L137
- Napolitano N R, Capaccioli M, Romanowsky A J, Douglas N G, Merrifield M R, Kuijken K, Arnaboldi M, Gerhard O and Freeman K C 2005 *Mon. Not. Roy. Ast. Soc.* **357** 691–706
- Navarro J F and White S D M 1994 *Mon. Not. Roy. Ast. Soc.* **267** 401–12
- Navarro J F, Frenk C S and White S D M 1997 *Astrophys. J.* **490** 493–508
- Newberg H J *et al* 2002 *Astrophys. J.* **569** 245–74
- Noordermeer E 2006 *The Distribution of Gas Stars and Dark Matter in early-type disk galaxies* PhD thesis, University of Groningen, The Netherlands
- Noordermeer E, van der Hulst J M, Sancisi R, Swaters R S and van Albada T S 2007 *Mon. Not. Roy. Ast. Soc.* **376** 1515–46
- Norris M A, Sharples R M and Kuntschner H 2006 *Mon. Not. Roy. Ast. Soc.* **367** 815–24
- Novati S C, de Luca F, Jetzer P and Scarpetta G 2006 *Astron. Astrophys.* **459** 407–14
- O’Connell R W 1980 *Astrophys. J.* **236** 430–40
- O’Dea C P, Baum S A, Mack J and Koekemoer A M 2004 *Astrophys. J.* **612** 131–51
- Odenkirchen M, Grebel E K, Dehnen W, Rix H-W, Yanny B, Newberg H J, Rockosi C M, Martínez-Delgado D, Brinkmann J and Pier J R 2003 *Astron. J.* **126** 2385–407
- Oey M S and Clarke C J 1997 *Mon. Not. Roy. Ast. Soc.* **289** 570–88
- Olling R P and Merrifield M R 1998 *Mon. Not. Roy. Ast. Soc.* **297** 943–52
- Olling R P and Merrifield M R 2001 *Mon. Not. Roy. Ast. Soc.* **326** 164–80
- O’Neil K, Bothun G, van Driel, W and Monnier Ragaigne D *et al* 2004 *Astron. Astrophys.* **428** 823–35
- Oort J 1932 *Bull. Astron. Inst. Netherlands* **6** 249–87
- Oort J A 1961 *The Distribution and Motion of Interstellar Matter in Galaxies* (New York: Benjamin), pp 3–12
- Oppenheimer B R, Hambly N C, Digby A P, Hodgkin S T and Saumon D 2001 *Science* **292** 698–702
- Palunas P and Williams T B 2000 *Astron. J.* **120** 2884–903
- Panther B, Heavens A F and Jimenez R 2003 *Mon. Not. Roy. Ast. Soc.* **343** 1145–54
- Pearce F R, Jenkins A, Frenk C S, White S D M, Thomas P A, Couchman H M P, Peacock J A and Efstathiou G 2001 *Mon. Not. Roy. Ast. Soc.* **326** 649–66
- Peletier R F 1989, PhD Thesis University of Groningen, The Netherlands
- Peletier R F, Davies R L, Illingworth G D, Davis L E and Cawson M 1990 *Astron. J.* **100** 1091–142
- Peletier R F, Balcells M, Davies R L, Andrekakis Y, Vazdekis A, Burkert A and Prada F 1999 *Mon. Not. Roy. Ast. Soc.* **310** 703–16
- Persic M, Salucci P and Stel F 1996 *Mon. Not. Roy. Ast. Soc.* **281** 27–47
- Peterson B M 2004 *Astrophys. J.* **613** 682–99
- Pfenniger D and Revaz Y 2005 *Astron. Astrophys.* **431** 511–6
- Pidopryhora Y, Lockman F J and Shields J C 2007 *Astrophys. J.* **656** 928–42

- Pierce M *et al* 2006 *Mon. Not. Roy. Ast. Soc.* **366** 1253–64
- Pizagno J, Prada F, Weinberg D H, Rix H-W, Harbeck D, Grebel E K, Bell E F, Brinkmann J, Holtzman J and West A 2005 *Astrophys. J.* **633** 844–56
- Poggianti B, Bridges T J, Komiyama Y, Yagi M, Carter D, Mobasher B, Okamura S and Kashikawa N 2004 *Astrophys. J.* **601** 197–213
- Press W H and Schechter P 1974 *Astrophys. J.* **187** 425–38
- Prochaska L C, Rose J A, Caldwell N, Castilho B, Concannon K, Harding P, Morrison H and Schiavon R P 2007 *Astron. J.* in press
- Proctor R N and Sansom A E 2002 *Mon. Not. Roy. Ast. Soc.* **333** 517–43
- Quillen A C, Brookes M H, Keene J, Stern D, Lawrence C R and Werner M W 2006 *Astrophys. J.* **645** 1092–1101
- Rafferty D A, McNamara B R, Nulsen P E J and Wise M W 2006 *Astrophys. J.* **652** 216–31
- Reddy B E, Lambert D L. and Prieto C A 2006 *Mon. Not. Roy. Ast. Soc.* **367** 1329–66
- Reid I N, Sahu K C and Hawley S L 2001 *Astrophys. J.* **559** 942–7
- Reid I N 2005 *Ann. Rev. Astron. Astrophys.* **43** 247–92
- Renzini A 1986, *Stellar Populations* (Cambridge: Cambridge University Press) p 213–23
- Revnivtsev M G, Churazov E M, Sazonov S Yu, Sunyaev R A, Lutovinov A A, Gilfanov M R, Vikhlinin A A, Shtykovsky P E and Pavlinsky M N 2004 *Astron. Astrophys.* **425** L49–52
- Rodgers A W and Paltoglou G 1984 *Astrophys. J.* **283** L5–7
- Rogstad D H, Wright M and Lockhard I 1976 *Astrophys. J.* **204** 703–11
- Rogstad D H and Shostak G S 1972 *Astrophys. J.* **176** 315–21
- Romanowsky A J, Douglas N, Arnaboldi M, Kuijken K, Merrifield M R, Napolitano N R, Capaccioli M and Freeman K C 2003 *Science* **301** 1696–8
- Rose J A, Arimoto N, Caldwell C N, Schiavon R P, Vazdekis A and Yamada Y 2005 *Astron. J.* **129** 712–28
- Rubin V C, Thonnard N and Ford W K 1978 *Astrophys. J.* **225** L107–11
- Rusin D and Kochanek C S 2005 *Astrophys. J.* **623** 666–82
- Rusin V C, Keeton C R and Winn J N 2005 *Astrophys. J.* **627** L95–6
- Ryden B 2002 *Introduction to Cosmology* (New York: Addison-Wesley)
- Sackett P D 1997 *Astrophys. J.* **483** 103–10
- Salim S *et al* 2005 *Astrophys. J.* **619** L39–42
- Salpeter E E 1955 *Astrophys. J.* **121** 161–7
- Sanchez-Blazquez P, Gorgas J, Cardiel N and Gonzalez J J 2006 *Astron. Astrophys.* **457** 809–21
- Sanders R H and McGaugh S S 2002 *Ann. Rev. Astron. Astrophys.* **40** 263–317
- Sargent W L W, Young P J, Bokserberg A, Shortridge K, Lynds C R and Hartwick P D A 1978 *Astrophys. J.* **221** 731–44
- Saviane I, Held E V and Bertelli G 2000 *Astron. Astrophys.* **355** 56–68
- Schaye J 2004 *Astrophys. J.* **609** 667–82
- Schechter P L 1976 *Astrophys. J.* **203** 297–306
- Schechter P L and Gunn J E 1979 *Astrophys. J.* **229** 472–84
- Schiavon R P, Faber S M, Castilho B V and Rose J A 2002 *Astrophys. J.* **580** 850–72
- Schiavon R P, Faber S M, Rose J A and Castilho, B V 2002 *Astrophys. J.* **580** 873–86
- Schiavon R P, Caldwell C N and Rose J A 2004 *Astron. J.* **127** 1513–30
- Shlosman, I 2001 *The Central Kiloparsec of Starbursts and AGN: The La Palma Connection* (San Francisco: Astro. Soc. Pacific) *ASP Conf. Proc.* vol 249) pp 55–77
- Schmidt M 1959 *Astrophys. J.* **129** 243–58
- Schombert J M and Bothun G D 1987 *Astron. J.* **93** 60–73
- Schröder K P and Cuntz M 2005 *Astrophys. J.* **630** L73–6
- Schödel R, Ott T, Genzel R, Eckart A, Mouawad N and Alexander T 2003 *Astrophys. J.* **596** 1015–34
- Schulz S and Struck C 2001 *Mon. Not. Roy. Ast. Soc.* **328** 185–202
- Schwarzschild M 1958 *Structure and Evolution of the Stars* (Princeton: Princeton Univ. Press)
- Schwarzschild M 1979 *Astrophys. J.* **232** 236–47
- Searle L, Sargent W L W and Bagnuolo W G 1973 *Astrophys. J.* **179** 427–38
- Searle L and Zinn H 1978 *Astrophys. J.* **225** 357–79
- Seaton M J, Yan Y, Mihalas D and Pradhan A K 1994 *Mon. Not. Roy. Ast. Soc.* **266** 805–28
- Seaton 1996 *Mon. Not. Roy. Ast. Soc.* **279** 95–100
- Sellwood J A and Wilkinson A 1993 *Rep. Prog. Phys.* **56** 173–256
- Sérsic J L 1968 *Atlas de Galaxias Australes* (Cordoba Argentina: Observatorio Astronomico)
- Shetrone M, Venn K A, Tolstoy E, Primas F, Hill V and Kaufer A 2003 *Astron. J.* **125** 684–706
- Siegel M H, Majewski S R, Reid I N and Thompson I B 2002 *Astrophys. J.* **578** 151–75
- Silk J 1997 *Astrophys. J.* **481** 703–9
- Solanes J M, Manrique A, Garcia-Gomez C, Gonzalez-Casado G., Giovanelli R and Haynes M P 2001 *Astrophys. J.* **548** 97–113

- Somerville R S and Kolatt 1999 *Mon. Not. Roy. Ast. Soc.* **305** 1–14
- Somerville R S and Primack J R 1999 *Mon. Not. Roy. Ast. Soc.* **310** 1087–1110
- Spinrad H and Taylor B J 1971 *Astrophys. J. Supp.* **22** 445–84
- Springel V and Hernquist L 2003 *Mon. Not. Roy. Ast. Soc.* **339** 289–311
- Springel V, Di Matteo T and Hernquist L 2005 *Mon. Not. Roy. Ast. Soc.* **361** 776–794
- Squires G, Kaiser N, Babul A, Falhman G, Woods D, Neumann D and Böhringer H 1996 *Astrophys. J.* **461** 572–86
- Steinmetz M and Navarro J F 1999 *Astrophys. J.* **513** 555–60
- Statler T S, Emsellem E, Peletier R F and Bacon R 2004 *Mon. Not. Roy. Ast. Soc.* **353** 1–14
- Straniero O, Chieffi A and Limongi M 1997 *Astrophys. J.* **490** 425–36
- Strickland D K, Heckman T M, Colbert E J M, Hoopes C G and Weaver K A 2004 *Astrophys. J. Supp.* **151** 193–236
- Strutskie M F *et al* 2006 *Astron. J.* **131** 1163–83
- Swaters R A, Sancisi R and van der Hulst J M 1997 *Astrophys. J.* **366** 449–66
- Thilker D *et al* 2005 *Astrophys. J.* **619** L79–92
- Thomas D, Maraston S and Bender R 2003 *Mon. Not. Roy. Ast. Soc.* **339** 897–911
- Tonry J and Schneider D P 1988 *Astron. J.* **96** 807–15
- Toomre A 1981 *Structure and Evolution of Normal Galaxies* (Cambridge: Cambridge University Press) pp 111–36
- Toomre A 1964 *Astrophys. J.* **139** 1217–38
- Trager S, Faber S M, Worthey G and González J J 2000 *Astron. J.* **119** 1645–76
- Trager S 2004 *Origin and Evolution of the Elements* (Carnegie Observatories Astrophys Series vol 4) ed A McWilliam and M Rauch (Cambridge: Cambridge Univ Press) p 391
- Tremaine S *et al* 2002 *Astrophys. J.* **574** 740–53
- Tripicco M and Bell R A 1995 *Astron. J.* **110** 3035–49
- Trujillo L, Graham A W and Caon N 2001 *Mon. Not. Roy. Ast. Soc.* **326** 869–76
- Tsikoudi V 1979 *Astrophys. J.* **234** 842–53
- Tsujimoto T and Shigeyama T 1998 *Astrophys. J.* **508** L151–4
- di Tullio G A 1979 *Astron. Astrophys.* **37** 591–600
- Tully R B and Fisher J R 1977 *Astron. Astrophys.* **54** 661–73
- Tully R B and Mohayaee R 2004 *Outskirts of Galaxy Clusters: Intense Life in the Suburbs* ed A Diaferio IAU Colloquium 195 p 205–11 Torino Italy (Cambridge: Cambridge University Press) (Preprint astro-ph/0404006)
- Udalski A, Szymanski M, Stanek K Z, Kaluzny J, Kubiak M, Mateo M, Krzeminski W, Paczynski B and Venkat R 1994 *Acta Astron.* **44** 165–89
- van Albada T S and Sancisi R 1986 *Royal Soc. of London Philosophical Transactions Series A* **320** 447–64
- van Dokkum P G, Franx M, Fabricant D, Kelson D D and Illingworth G D 1999 *Astrophys. J.* **520** L95–8
- Vader J P, Vigroux L, Lachize-Rey M and Souviron J 1988 *Astron. Astrophys.* **203**, 217–25
- Valdes F, Gupta R, Rose J A, Singh H P and Bell D J 2004 *Astrophys. J. Supp.* **152** 251–9
- Valluri M and Merritt D 1998 *Astrophys. J.* **506** 686–711
- Vazdekis A, Casuso E, Peletier R F and Beckman J E 1996 *Astrophys. J. Supp.* **106** 307–39
- Veilleux S, Cecil G and Bland-Hawthorn J 2005 *Ann. Rev. Astron. Astrophys.* **43** 769–826
- Venn K A, Irwin, M, Shetrone M D, Tout C A, Hill V and Tolstoy E 2004 *Astron. J.* **128** 1177–95
- van de Ven G, Hunter C, Verolme E K and de Zeeuw P T 2003 *Mon. Not. Roy. Ast. Soc.* **342** 1056–82
- van Zee L, Skillman E D and Haynes M P 2004 *Astron. J.* **128** 121–36
- Verheijen M A W 2001 *Astrophys. J.* **563** 694–715
- Verolme E, Cappellari M, Copin Y, van der Marel R P, Bacon R, Bureau M, Davies R L, Miller B M and de Zeeuw P T 2002 *Mon. Not. Roy. Ast. Soc.* **335** 517–25
- Vollmer B, Soida M, Otmianowska-Mazur K, Kenney J D P, van Gorkom J H and Beck R 2006 *Astron. Astrophys.* **453** 883–93
- Volonteri M, Madau P and Haardt F 2003 *Astrophys. J.* **593** 661–6
- Wallerstein G *et al* 1997 *Rev Mod Phys* **69** 995–1084
- Wake D A *et al* 2006 *Mon. Not. Roy. Ast. Soc.* **372** 537–50
- Weinberg M D and Blitz L 2006 *Astrophys. J.* **641** L33–6
- Weiner B J, Sellwood J A and Williams T B 2001 *Astrophys. J.* **546** 931–51
- Weiss A, Serenelli A, Kitsikis A, Schlattl H and Christensen-Dalsgaard J 2005 *Astron. Astrophys.* **441** 1129–33
- White S D M and Rees M J 1978 *Mon. Not. Roy. Ast. Soc.* **183** 341–58
- Whitmore B C and Schweizer F 1995 *Astron. J.* **109** 960–80
- Whyte L F, Abraham R G, Merrifield M R, Eskridge P B, Frogel J A and Pogge R W 2002 *Mon. Not. Roy. Ast. Soc.* **336** 1281–6
- Willman B *et al* 2005 *Astrophys. J.* **626** L85–8
- Worthey G 1994 *Astrophys. J. Supp.* **95** 107–49
- Worthey, G 2004 *Astron. J.* **128** 2826–37
- Wu H, Shao Z, Mo H J, Xia X and Deng Z 2005 *Astrophys. J.* **622** 244–59
- Wyithe J S B and Loeb A 2003 *Astrophys. J.* **595** 614–23

- Yoachim P and Dalcanton J J 2006 *Astron. J.* **131** 226–49
- Yoon S-J, Yi S K and Lee Y-W 2006 *Science* **311** 1129–32
- York D G *et al* 2000 *Astron. J.* **120** 1579–87
- de Zeeuw P T *et al* 2002 *Mon. Not. Roy. Ast. Soc.* **329** 513–30
- Zhang J, Fall S M and Whitmore B C 2001 *Astrophys. J.* **561** 727–50
- Zoccali M, Renzini A, Ortolani S, Greggio L, Saviane I, Cassisi S, Rejkuba M, Barbuy B, Rich R M and Bica E 2003 *Astron. Astrophys.* **399** 931–56
- Zwicky F 1937 *Astrophys. J.* **86** 217–46